

AD-A049 564

HONEYWELL INC MINNEAPOLIS MINN SYSTEMS AND RESEARCH --ETC F/G 17/5
AUTOMATED IMAGE ENHANCEMENT TECHNIQUES FOR SECOND GENERATION FL--ETC(U)
DEC 77

DAA053-76-C-0195

NL

UNCLASSIFIED

77SRC93

AD
A049564

OF 4



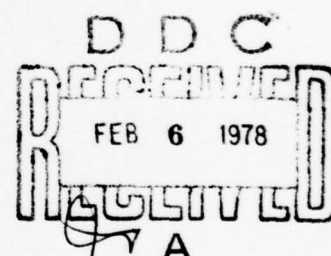


AD A 049564

JDC FILE COPY



Honeywell



FINAL REPORT
ON AUTOMATED IMAGE
ENHANCEMENT TECHNIQUES
FOR SECOND GENERATION
FLIR

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

ACCESSION FOR	Write Section <input checked="" type="checkbox"/>	Unit Section <input type="checkbox"/>
NTIS	UNANNOUNCED	JUSTIFIED
DOC	BY DISTRIBUTION AVAILABILITY CODES	
BY ATAIL DOC OF SPECIAL		
DUAL		

with original

A

2

Honeywell

14

77SRC93

11

December 1977

9

12 354 p.

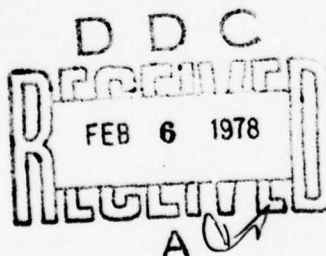
6

FINAL REPORT ON AUTOMATED IMAGE ENHANCEMENT TECHNIQUES FOR SECOND GENERATION FLIR

15

Contract No. DAAG53-76-C-0195

Prepared for
U.S. Army Electronics Command
DELNV-VI, Night Vision & Electro-Optical Laboratories
Fort Belvoir, Virginia 22060



DISTRIBUTION STATEMENT A
Approved for public release,
Distribution Unlimited

Systems & Research Center

2700 RIDGWAY PARKWAY
MINNEAPOLIS, MINNESOTA 55413

Printed in U.S.A.

402349

B

ABSTRACT

PART I

Image enhancement algorithms were studied for thermal imager contrast, signal-to-noise ratio (sensitivity), and resolution enhancement; detector responsivity equalization; and DC restoration.

CONTRAST ENHANCEMENT

Several schemes were investigated for enhancing local contrast of FLIR imagers. An efficient, real time, implementable algorithm for local area gain brightness control was developed to provide total hands-off gain and bias control of the FLIR. This scheme not only completely automates the manual gain and bias controls, but also modifies the local area gain and bias to make optimal use of the narrow dynamic range of the display. This in turn results in crisper thermal imagery and improved local contrast. This simple and efficient recursive real time charge-coupled device (CCD) implementation uses just two CCD line delays.

S/N ENHANCEMENT

Several schemes for within-frame image noise smoothing were investigated. Promising adaptive algorithms and their hardware implementations--such as the two-dimensional separable median filter and curvature-directed adaptive filters--were developed. Also, a simple and effective real time implementable scheme was investigated for registering and averaging

successive frames. This smooths the image noise to improve MRT by averaging in time. A straightforward hardware design was developed to accomplish this function.

RESOLUTION ENHANCEMENT

Resolution restoration schemes were investigated to restore the optics blur both up to and even beyond the Rayleigh diffraction limit. These schemes were shown to be applicable only when the FLIR imagers have adequate sampling (at better than the Nyquist rate). Implementations were designed for the linear Wiener filter and the stochastic approximation superresolution algorithms.

RESPONSIVITY EQUALIZATION AND DC RESTORATION

Both deterministic and statistical schemes for detector responsivity equalization were developed and simulated. These schemes were shown to successfully equalize the detector bias and responsivity differences in a multidetector (either parallel scan or staring array) configuration. The related problem of DC restoration in parallel scan, AC-coupled thermal imagers was also investigated. A scheme for synthetic DC restoration was developed, simulated, and shown to be very effective.

A Monte Carlo approach to predicting the MRT curve of an image-enhanced FLIR was developed, tested and shown to be equivalent to the NVL computer model when all the system components are linear. This

approach has the advantage of being able to successfully incorporate non-linear and position variant effects (as with image enhancement).

Preliminary hardware implementation designs were made for the more promising of the image enhancement schemes. The emphasis was of course on real time implementation. Both on- and off-focal plane implementations were considered. These designs are modular and conform with the current trend toward modularity in military thermal imagers.

Several hundred thermal images were enhanced by computer simulation of each of the above image enhancement algorithms, singly, and in various combinations (cascade). The resultant images were analyzed using the statistical imagery features developed in this program to quantify the effect of each enhancement process on image contrast, target shape, texture, and noise.

PART II

The objective of this experimental evaluation was to discover whether observer performance was affected by changes in image quality caused by various enhancement techniques. A number of thermal imagers, views of the ground containing one or more military vehicles, were transformed using various enhancement algorithms. Observers were asked whether particular hot spots on the resultant images were tanks, armored personnel carriers, trucks, or jeeps. A total of 109 observers took part in the experiment. Using the Mann-Whitney, two-sample, two-tailed U test, the proportion of correct responses (accuracy) and the response

times for images transformed by the enhancement algorithms were compared with the data obtained with the untransformed, original images.

Most of the enhancement algorithms had little effect on response time although there were some significant changes: (1) there was a small improvement when noise-free images were treated with a combination of contrast and minimum resolvable temperature algorithms; (2) there was a decrement (i.e., longer response time) when a combination of contrast, minimum resolvable temperature, and resolution restoration algorithms was used on noisy images; and (3) there were also decrements when two, different minimum resolvable temperature algorithms were used on images with very large targets. With the accuracy comparisons, there were no significant differences between the transformed and the original images.

PART III

We developed a prediction model of visual search which included the effects of target and background features. The model was based on first partitioning the background into a number of homogeneous regions and then predicting performance within each region.

We performed an experiment to study the relationship between a broad range of objective and subjective measures of target and background characteristics and target acquisition performance. Subjects searched for vehicles and similar objects embedded by computer in a variety of scenes.

The three objective measures most related to performance were luminance contrast, texture contrast, and display resolution. However, performance was even more strongly related to judged difficulty of locating and recognizing the target together with display resolution.

CONTENTS--PART I

Section		Page
I	INTRODUCTION	1
	Need for Image Enhancement	1
	Summary of Part I	3
	Report Organization	5
II	DC RESTORATION AND DETECTOR RESPONSIVITY EQUALIZATION	7
	DC Restoration	7
	A Synthetic DC Restore Algorithm	13
	DC Restore Performance Evaluation	16
	Is DC Restoration Necessary in Second Generation FLIRs?	21
	Detector Responsivity Equalization	22
III	STATISTICAL CHARACTERIZATION OF FLIR IMAGERY	24
	Shape Statistics	26
	Intensity and Contrast Statistics	33
	Texture Features	41
	Grey-Level Difference Statistics	41
	Interpretation of the Texture Features	45

CONTENTS--PART I (continued)

Section	Page
IV CONTRAST ENHANCEMENT	53
Introduction	53
The Need for Contrast Enhancement	53
High Frequency Emphasis Spatial Filters	55
Local Area Gain/Brightness Control	62
Statistical Analysis of Contrast Enhanced Imagery	74
Contrast and Intensity Statistics	75
Texture	86
Contrast Enhancement Summary	87
V MRT ENHANCEMENT	92
Introduction	92
Intraframe Smoothing Filters	94
Median Filters	94
Separable Median Filter	95
Adaptive Variable Width Filters	101
Statistical Analysis of Intraframe Smoothing Algorithms	107
Intensity and Contrast Statistics	108
Texture Features	109
Intraframe MRT Enhancement Summary	109
Interframe Noise Smoothing	114
Summary	122

CONTENTS--PART I (continued)

Section	Page
VI RESOLUTION ENHANCEMENT	124
Optics and Detector Model	126
Inverse Filtering for Full Frame Resolution Restoration	131
Summary and Conclusion of the Full Frame Resolution Restoration Task	132
Superresolution Algorithms	134
Digital Image Magnification	135
Summary and Conclusions on Superresolution	136
VII INTEGRATED IMAGE ENHANCEMENT (CONTRAST, MRT, RESOLUTION)	139
Cascade Processing	139
Integrated Approach	145
Two-Channel Image Enhancement Algorithms (by Thomas S. Huang)	145
Summary and Conclusions	149
VIII IMAGE ENHANCED FLIR PERFORMANCE MODEL	155
The Methodology	157
Generation of Bar Pattern	160
The Optics and Detector Blurring	160
Detector Noise	163
Electronic Transfer Function	164

CONTENTS--PART I (continued)

Section		Page
	Image Enhancement	165
	Display Transfer Function	167
	Matched Filter	167
	Measurement of Signal and Noise at the Output	167
	Iteration	169
	Scaling the Axes	170
	An Example	172
	Methodology for Prediction of Recognition Probability	176
	The Procedure	177
IX	SECOND GENERATION FLIR EXAMPLE	182
	Mission Parameters Specification	182
	Second Generation FLIR Design	185
X	IMPLEMENTATION OF IMAGE ENHANCEMENT ALGORITHMS	192
	Systems Requirements and Technology Considerations	193
	Introduction	193
	Video Delay and Storage	194
	Technology Considerations	195
	Focal Plane Partitioning	202
	Hardware Subelements	204

CONTENTS--PART I (concluded)

Section	Page
Image Enhancement Circuit Designs	207
Algorithm Selection	207

CONTENTS--PART II

Section		Page
	ACKNOWLEDGEMENTS	227
I	INTRODUCTION	228
II	METHOD	230
	Imagery	230
	Design and Experimental Conditions	234
	Apparatus and Procedure	238
	Observers	239
III	RESULTS	240
	Accuracy	240
	Response Time	247
IV	DISCUSSION	255
	Conclusions	256
	Comments	257

CONTENTS--PART III

Section		Page
I	INTRODUCTION	258
	Objectives	258
	Background	258
	The Search Model	259
	The Mathematical Structure of Search for a Target in a Homogeneous Field - Initial Formulation	261
	Present Formulation	265
II	METHOD	267
	Design	267
	Stimuli	268
	Difficulties	272
	Procedures	272
	Training	273
	Main Trials	274
III	DATA ANALYSIS	276
	Overview	276
	Analysis of the Imagery	276
	Target Measures	277
	Background Measures	278

CONTENTS--PART III (concluded)

Section	Page
III	
Measures on the Search Field--Target in Background	279
Psychometrics	282
Subject Performance Data	283
Parameter Estimation--Curve Fitting	284
Results	285
An Initial Comparison of Target Location and Identification	286
Predicting Target Location	286
Predicting Target Identification	290
Comments	292
REFERENCES	294
APPENDIX A	297
APPENDIX B	315
APPENDIX C	319

LIST OF ILLUSTRATIONS

Figure		Page
1	Rect Function Response of an RC Circuit	9
2	Effect of AC Coupling Degradation on a Typical DC-Restored FLIR Image	11
3	Synthetic DC Restoration of Test Target No. 1	12
4	Line-to-Line Pixel Differences from AC Coupling	14
5	Functional Block Diagram of the Synthetic DC Restore Algorithm	15
6	Circuit Response to a Rect Function	18
7	NVLI/6--Tank/(10 x 20/137)	27
8	Mean Slope of Edges vs. Perimeter/ Area	29
9	Mean vs. Standard Deviation of Relative Slopes	30
10	Mean vs. Standard Deviation of Edge Lengths	31
11	Target Intensity Histogram	34
12	Background Intensity Histogram	34
13	Mean and Standard Deviation of Intensities of Targets (T) and Backgrounds (B)	35
14	Skewness vs. Excess of Target (T) and Background (B) Intensities	36
15	Mean vs. Skewness of Target (T) and Background (B) Intensities	38
16	Illustrating Point Pairs for Grey-Level Difference Statistics	42

LIST OF ILLUSTRATIONS (continued)

Figure		Page
17	Grey-Level Difference Histograms for a Representative Target and Its Background for Various Values of δ	43-44
18	"Mean" vs. "Contrast" Texture Features for Spacing $\Delta = 4$	47
19	Mean Intensity vs. "Mean" Texture Feature for Spacing $\Delta = 1$, Background (B), Targets (T, A, J)	48
20	Mean Intensity vs. "Mean" Texture Feature for Spacing $\Delta = 2$	49
21	Mean Intensity vs. "Mean" Texture Feature for Spacing $\Delta = 4$	50
22	Mean Intensity vs. "Mean" Texture Feature for Spacing $\Delta = 8$	51
23(a)	High Emphasis Filter Using a Basic Low Pass Filter	57
23(b)	Desired High Frequency Response	57
24(a)	Original Thermal Image	59
24(b)	High Frequency Emphasis Nonrecursive Filter	59
24(c)	High Frequency Emphasis Recursive Filter	59
25(a)	Original Thermal Image	61
25(b)	High Frequency Emphasis Nonrecursive Filter	61
25(c)	High Frequency Emphasis Recursive Filter	61
26	A Recursive Filter Approach to Local Area Gain/Brightness Control	64

LIST OF ILLUSTRATIONS (continued)

Figure		Page
27(a)	Nonrecursive Local Area Gain/Brightness Control Transformed Version of Figure 24(a)	66
27(b)	Recursive Local Area Gain/Brightness Control Transformed Version of Figure 24(a)	66
28(a)	Nonrecursive Local Area Gain/Brightness Control Transformed Version of Figure 25(a)	67
28(b)	Recursive Local Area Gain/Brightness Control Transformed Version of Figure 25(a)	67
29(a)	Original Thermal Image	68
29(b)	Recursive LAGBC Enhanced	68
30(a)	Original Thermal Image	69
30(b)	Recursive LAGBC Enhanced	69
31(a)	Original Thermal Image	70
31(b)	Recursive LAGBC Enhanced	70
32(a)	Original Thermal Image	71
32(b)	Recursive LAGBC Enhanced	71
33	Histogram Equalized Version of the LAGBC Image in Figure 27(a)	73
34(a)	Mean Intensity of Target (T) and Background Before vs. After High Frequency Emphasis (Recursive) Filtering	76
34(b)	Mean Intensities of Target (T) and Background Before and After Local Area Gain/Brightness Control (Nonrecursive)	77

LIST OF ILLUSTRATIONS (continued)

Figure		Page
34(c)	Mean Intensity of Target (T) and Background Before vs. After Local Area Gain/Brightness Control (Recursive)	78
35(a)	Standard Deviation of Target (T) and Background (B) Intensities Before vs. After High Frequency Emphasis Filtering (Recursive)	80
35(b)	Local Area Gain/Brightness Control (Nonrecursive)	81
35(c)	Local Area Gain/Brightness Control (Recursive)	82
36(a)	Percentage Change in Peak Contrast After Enhancement vs. Original Peak Contrast	83
36(b)	Percentage Change in Average Contrast After Enhancement vs. Original Average Contrast	84
36(c)	Percentage Change in Edge Contrast After Enhancement vs. Original Edge Contrast	85
37(a)	"Mean" Texture Feature ($\Delta = 1$) Before vs. After LAGBC (Nonrecursive)	89
37(b)	"Mean" Texture Feature ($\Delta = 2$) Before vs. After LAGBC (Nonrecursive)	90
37(c)	"Mean" Texture Feature ($\Delta = 4$) Before vs. After LAGBC (Nonrecursive)	91
38(a)	Original MRT Curve and Intraframe Smoothed Curve	93
38(b)	Original and Interframe Smoothed MRT Curves	93
39	A 3 x 3 Median Filter	95
40	Separable Median Filter Concept	96

LIST OF ILLUSTRATIONS (continued)

Figure		Page
41	Separable Median Filter Structure	97
42(a)	Thermal Image with Noise Added, SNR=3	99
42(b)	5 x 5 Median Filtered	99
42(c)	5 x 5 Separable Median	99
43(a)	Thermal Image with Noise Added, SNR=3	100
43(b)	Median Filtered 3 x 3 S/N Gain 1.0 dB	100
43(c)	Median Filtered 5 x 5 S/N Gain = 4.3 dB	100
44	Adaptive Variable Width Intraframe Smoothing	102
45(a)	Recursive Adaptive Filtered Version of the Noisy Image in Figure 42(a) (S/N Gain = 2.13 dB)	104
45(b)	Nonrecursive Adaptive Filtered Version of the Noisy Image in Figure 42(a) (S/N Gain = 3.86 dB)	104
46(a)	Mean Intensity of Targets (T) and Backgrounds (B) Before vs. After 5 x 5 Median Filtering	105
46(b)	Standard Deviation of Intensity of Targets (T) and Backgrounds (B) Before and After 5 x 5 Median Filtering	106
47(a)	Percentage Change in Peak Contrast vs. Original Peak Contrast	110
47(b)	Percentage Change in Average Contrast vs. Original Average Contrast	110
47(c)	Percentage Change in Edge Contrast vs. Original Edge Contrast	110

LIST OF ILLUSTRATIONS (continued)

Figure		Page
48(a)	"Mean" Texture Feature Before vs. After 5 x 5 Median Filtering (Spacing $\Delta = 1$)	111
48(b)	"Mean" Texture Feature Before vs. After 5 x 5 Median Filtering (Spacing $\Delta = 2$)	112
48(c)	"Mean" Texture Feature Before vs. After 5 x 5 Median Filtering (Spacing $\Delta = 4$)	113
49(a)	Smoothed Reference Window	116
49(b)	Smoothed Correlation Window Used in the Correlation Tracker Algorithm for Translational Registration	116
50(a)	First Noisy Frame SNR = 3	118
50(b)	After Averaging Two Noisy Frames	118
50(c)	After Averaging Four Noisy Frames	118
50(d)	After Averaging Nine Noisy Frames	118
51	Signal-to-Noise Ratio Gains Acruued by Frame Averaging vs. the Number of Frames Stacked for the Stationary Target Case	119
52	Two Successive Video Frames Brought into Registration by Shift/Rotation/Stretching	121
53	Detector and Optics Blur (Gaussian) Proportions Assumed	128
54(a)	Combined Optics and Detector MTF	129
54(b)	Wiener Inverse Filter Response	129

LIST OF ILLUSTRATIONS (continued)

Figure		Page
55(a)	FLIR Image Blurred with a Space Variant MTF Shown in Figure 22(a) (Five Pixels Blur on Axis and Seven Off Axis)	133
55(b)	Restored by Piecewise Space Invariant Wiener Filter	133
56(a)	Subframe of a FLIR Image Containing a 1-1/4 Ton Truck	137
56(b)	Digitally Magnified Version of Figure 56(a)	137
56(c)	SA Algorithm Applied to the Magnified Image in Figure 56(b)	137
57(a)	LAGBC Only	141
57(b)	LAGBC + Med	141
57(c)	3 x 3 Med + LAGBC	141
57(d)	LAGBC + 5 x 5 Med	141
57(e)	5 x 5 Med + LAGBC	142
57(f)	LAGBC + Recursive Adaptive Smoothing Filter	142
57(g)	Recursive Adaptive Smoothing Filter + LAGBC	142
58	Flow Chart of Cascade Processing for Human Factors Evaluation	144
59(a)	Noisy Thermal Image	146
59(b)	LAGBC + MRT	146
59(c)	Digital Magnification	146

LIST OF ILLUSTRATIONS (continued)

Figure		Page
59(d)	SA Resolution Restoration	146
60	Enhancement Algorithm 1	148
61	Enhancement Algorithm 2	148
62(a)	Original FLIR Image	150
62(b)	Output of Adaptive Filter	150
62(c)	Edge Regions Where the Median Filter Is Operative	150
62(d)	Final Output of Algorithm 2	150
63(a)	Output of 3 x 3 Low Pass Filter	151
63(b)	Band Pass Filter before Thresholding	151
63(c)	Band Pass Filter after Thresholding	151
63(d)	Final Output (Algorithm 1)	151
64(a)	Original	152
64(b)	Integrated Algorithm 2	152
64(c)	Integrated Algorithm 1	152
64(d)	Cascaded Result (LAGBC + MRT) for Comparison	152
65(a)	Original	153
65(b)	Integrated Algorithm 2	153
65(c)	Integrated Algorithm 1	153
65(d)	Cascaded Result (LAGBC + MRT) for Comparison	153

LIST OF ILLUSTRATIONS (continued)

Figure		Page
66	Image Enhanced FLIR Model	158
67	Simulation Outline for Predicting the MRT	159
68	Detector and Optics (Gaussian) Proportions Assumed	162
69	Measurement of Output Noise	166
70	Predicted MRT Curves Using the Monte Carlo Approach f_o = detector cutoff	173
71	Focal Plane Concept	186
72	High Performance FLIR Block Diagram	187
73	Predicted FLIR System MRT Using NVL Programs	190
74	CCD Split-Electrode Transversal Filter	197
75	N-Stage Video Delay with CTDs	200
76	Absolute Value Scheme with CCDs	203
77	Two-dimensional Nonrecursive, Nonseparable Filter	205
78	Two-dimensional Nonrecursive, Separable Filter	206
79	Two-dimensional Filter, One-dimensional Non-recursive, One-dimensional Recursive	208
80	Two-dimensional Recursive, Separable Filter	209
81	Recursive Low Pass Filter Implementation	211
82	Recursive Approach to Local Area Gain/Brightness Control	212

LIST OF ILLUSTRATIONS (concluded)

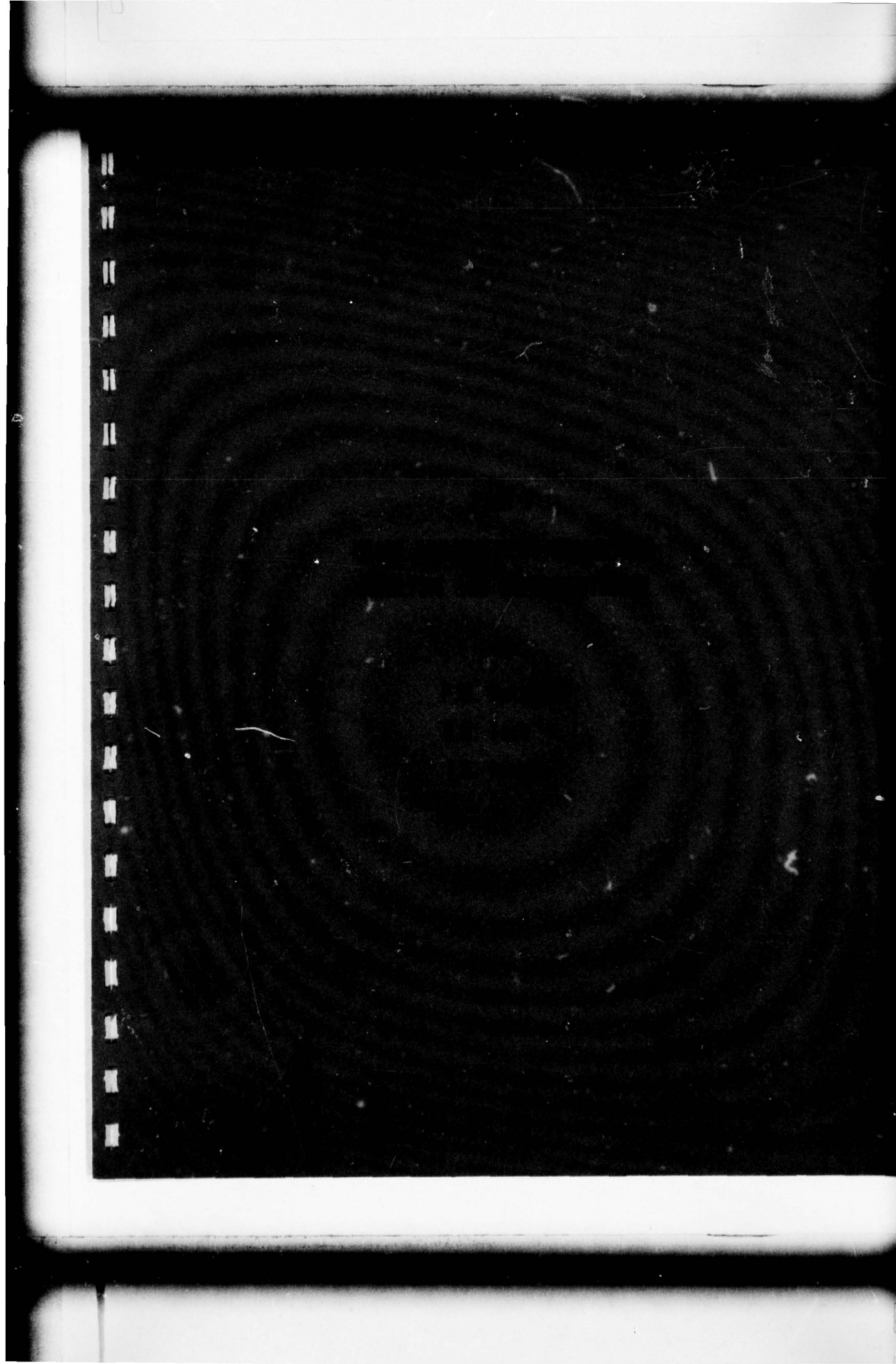
Figure		Page
83	Three-Pixel Median Filter	214
84	Separable Median Filter Structure	216
85	Adaptive Intraframe Smoothing	217
86	Local Curvature-Directed Adaptive Filtering Implementation	219
87	Squared Canonical Product Correlation for Frame Registration	221
88	SCPC Hardware Implementation	222
89	Superresolution Magnification and Interpolation	225
90	Stochastic Approximation Implementation	226
91	Homogeneous Regions	262
92	Graph of Equation (14)	265
93	Graph of Equation (15)	266
94	Main Features of the Experimental Design	268
95	Four Scenes	269
96	Six Targets	271
97	Modified Clock	274
98	Probability of Locating Target as a Function of Luminance Contrast and Texture Contrast	288
99	Judged Difficulty of Locating Target	289

LIST OF TABLES

Table		Page
1	Shape Statistics Summary	32
2	Summary of Contrast Statistics before Enhancement	40
3	Probability of Recognition Predicted by Model	179
4	Measured Probabilities of Recognition from the Evaluation Phase	180
5	FLIR System Requirements	184
6	FLIR System Parameters	191
7	Video Storage and Delay Requirements for One Pixel, One Line, One Field	194
8	Details of Images Used in Experiment	232-233
9	Experimental Conditions	235
10	Treatment Conditions Used with Each Subset of Images	236
11	Number of Correct Responses (a), Total Number of Trials (b), and Proportion of Correct Responses (c) for Each Scene by Treatment Combination	241-246
12	Total Number of Trials and Numbers and Proportions of Correct Responses, Incorrect Responses and Non-responses for Each Group of Scenes in Each Treatment Tested	248-249
13	Geometric Mean Times for Scene by Treatment Combinations, in Seconds	251-252
14	Arithmetic Mean, Geometric Mean, and Median Times for Each Group of Scenes in Each Treatment Tested, in Seconds	253-254

LIST OF TABLES (concluded)

Table		Page
15	Target Measures	277
16	Background Measures	278
17	Correlation between a Linear Combination of Predictor Variables and Performance Criteria	291



SECTION I

INTRODUCTION

This is Part I of the Final Report for Contract DAAG-53-76-C-0195, "Automatic Image Enhancement Techniques for Second Generation FLIR."

✓ The report presents the results of Phase I and Phase II studies in the application of image processing techniques to FLIR imaging systems. Part I describes the image enhancement algorithms investigated, their analysis, and preliminary design of the hardware to implement these functions. Part II presents the results of the human factors enhancement evaluation study, conducted as an adjunct to the above contract, in which the algorithms developed during this study were evaluated. Part III presents the development of a search effectiveness model for predicting search performance using electrooptical sensors in tactical situations. ↗

NEED FOR IMAGE ENHANCEMENT

An excellent case in favor of image enhancement techniques for night vision imagers was made by Mr. John Dehne of NVL.* He identified the three major areas in current generation FLIRs (forward looking infrared sensors) that would benefit from image enhancement. The first area, contrast enhancement, refers to the enhancement of local scene contrast and the automation of the gain and bias controls on FLIRs in a totally

*J. A. Dehne, "Application of Adaptive Image Processing Techniques to Night Vision Problems," Workshop on Application of Interactive Cybernetics Systems, October 13-16, 1975.

hands-off mode. This enables the varying and large scene temperature ranges to be squeezed adaptively into the limited display dynamic range without sacrificing the local target/background contrast. Automation of these controls frees the FLIR operator to perform more important target acquisition functions.

The second area involves the improvement of system sensitivity, i.e., signal-to-noise ratio. Minimum resolvable temperature (MRT) enhancement is needed with low temperature contrast scenes (under adverse weather conditions, for example) in which the detector and background noise dominate.

Resolution enhancement refers to the correction of the blur caused by the optics and the detector. Resolution enhancement is of importance when small aperture optics are dictated because of size considerations. Resolution enhancement then provides a way to improve the system resolution in spite of the small size optics.

Other areas that can benefit from image enhancement include specific defects associated with certain classes of current generation FLIRS, such as the DC restoration problem in AC coupled parallel scan devices. Responsivity equalization of detectors in a multidetector configuration is another related problem in search of a signal/image processing solution.

The above problems were addressed in this study of image processing techniques to improve second generation FLIR performance. The development, analysis, and implementation of algorithms for performing these functions form the body of this report.

SUMMARY OF PART I

Several schemes were investigated for enhancing local contrast of FLIR imagers. An efficient, real time implementable algorithm for local area gain brightness control was developed to provide total hands-off gain and bias control of the FLIR. This scheme not only completely automates the manual gain and bias controls, but also adaptively modifies the local area gain and bias to make optimal use of the narrow dynamic range of the display. This in turn results in crisper thermal imagery and improved local contrast. This simple and efficient recursive real time charge-coupled device (CCD) implementation uses just two CCD line delays. The hardware design that resulted from this study has been breadboarded (with internal funds) and is operational at the time of this writing.

Several schemes for within-frame image noise smoothing were investigated. Promising adaptive algorithms and their hardware implementations--such as the two-dimensional separable median filter and curvature-directed adaptive filters--were developed. Also, a simple and effective real time implementable scheme was investigated for registering and averaging successive frames. This smooths the image noise to improve MRT by averaging in time. A straightforward hardware design was developed to accomplish this function.

Resolution restoration schemes were investigated to restore the optics blur both up to and even beyond the Rayleigh diffraction limit. These schemes were shown to be applicable only when the FLIR imagers have adequate sampling (at better than the Nyquist rate). Current detector-limited FLIRs do not satisfy this criterion. Real time and near-real time

implementations were designed for the linear Wiener filter and the stochastic approximation super resolution algorithms.

Both deterministic and statistical schemes for detector responsivity equalization were developed and simulated. These schemes were shown to successfully equalize the detector bias and responsivity differences in a multidetector (either parallel scan or staring array) configuration. The related problem of DC restoration in parallel scan, AC coupled thermal imagers was also investigated. A scheme for synthetic DC restoration was developed, simulated, and found to be very effective.

A Monte Carlo approach to predicting the MRT curve of an image enhanced FLIR was developed, tested, and found to be equivalent to the NVL computer model when all the system components are linear. This approach also has the advantage of being able to successfully incorporate nonlinear and position variant effects (as with image enhancement). Therefore, it promises to be a powerful tool in image enhanced FLIR system design and evaluation.

Preliminary hardware implementation designs were made for the more promising of the above image enhancement schemes. The emphasis was of course on real time implementation. Both on- and off-focal plane designs were considered. Basic functions, such as background subtraction and antiblooming, can be incorporated on the focal plane. But most of the image enhancement algorithms were found to be more easily realizable off the focal plane on the multiplexed serial video stream. This characteristic lends the advantage of being independent of the specific focal plane structure of the FLIR. Therefore, these

designs are modular and conform with the current trend toward modularity in military thermal imagers.

Several hundred thermal images were enhanced by computer simulation of each of the above image enhancement algorithms, singly, and in various combinations (cascade). The resultant images were analyzed using the statistical imagery features developed in this program to quantify the effect of each enhancement process on image contrast, target shape, texture, and noise.

Over 400 enhanced thermal images were evaluated in a human factors evaluation study that measured their effect on recognition probability and response time. The results of this study are reported in Part II of this report.

REPORT ORGANIZATION

The following headings, which correspond to the major tasks of this program, form the outline of Part I:

- DC restoration and detector responsivity equalization
- Statistical characterization of FLIR imagery
- Contrast enhancement
- Minimum resolvable temperature (MRT) enhancement
- Resolution enhancement
- Integrated image enhancement (contrast, MRT, resolution)

- Image enhanced FLIR performance model
- Second generation FLIR example
- Implementation of image enhancement algorithms

This program was initiated by Mr. John Dehne, Mr. Peter Raimondi (technical monitor), Mr. Peter Van Atta (alternate technical monitor), and Mr. Thomas Cassidy (search effectiveness technical monitor) of NVL to study the application of image enhancement to the above problem areas. The Signal and Image Processing section and the Man-Machine Systems group at the Honeywell Systems and Research Center, and the Honeywell Radiation Center participated in this program. The program manager was Dr. M. Geokezas, Chief of the Signal and Image Processing section. Principal investigators were Dr. D. H. Tack and Dr. P. M. Narendra for image enhancement algorithm design and analysis; Dr. J. D. Joseph for image enhancement hardware design; Mr. J. Merchant for autofocus analysis; Dr. J. Bloomfield for the human factors image enhancement evaluation (Part II); and Dr. L. Williams for the search effectiveness model development (Part III).

SECTION II

DC RESTORATION AND DETECTOR RESPONSIVITY EQUALIZATION

This section discusses the DC restoration problem associated with AC coupled detector arrays, and detector gain and offset equalization in multi-detector focal planes.

The discussion of DC restoration is divided into two segments: 1) Analysis of current thermal imagers--serial and parallel scan, and the effect of AC coupling in these imagers with proposed solutions for synthetic DC restoration; and 2) summary of the discussion on current trends [3] in focal plane technology. This summary indicates that in monolithic and hybrid focal plane technology, DC coupling of the detectors to the focal plane CCD processors (multiplexers, for example) is becoming very feasible. This would make DC restoration unnecessary in second generation FLIRs.

The subsection on detector responsivity equalization summarizes the two approaches to gain and offset correction reported in the first two quarterly progress reports.

DC RESTORATION

DC restoration can be best illustrated with a current generation FLIR that employs AC coupling of the detectors to their electronics. We will consider a parallel scan FLIR (the Army Common Module FLIR) as the example

because of its ubiquity. In addition, the degradation due to AC coupling is much more severe in parallel scan FLIRs than in serial scan FLIRs. Following is an example of a current generation FLIR (Common Module FLIR).

The Common Module FLIR has a parallel scan geometry with (a maximum of) 180 detectors AC coupled to their amplifiers, which in turn drive a set of 180 LEDs that paint the image.

The practice of AC coupling arises for three reasons:

1. The detectors are photo-conductive and AC coupling eases the biasing considerations on the detector;
2. There is a need to limit the $1/f$ noise in the detectors;
3. The elimination of DC by AC coupling subtracts the average background component of the line from the video and thus increases the contrast sensitivity of the displayed image.

However, AC coupling of the detectors is also accompanied by some degradations on the resultant displayed image. The following paragraphs briefly review the AC coupling degradation problem in the context of the parallel scan FLIRs.

AC coupling degradations can be characterized by two essentially separate phenomena:

- Transient effects--undershoot at a rapid transition of temperature.
- Steady-state effects--loss of line-to-line correlation because of the loss of the average value of the line (streaking and droop).

The transient effect is shown in Figure 1, where the Rect function suffers a droop and an equal undershoot. As later analysis shows, the transient droop and undershoot problems are minor for the long time constants of the RC coupling circuitry in the Common Module FLIR and do not need to be corrected.

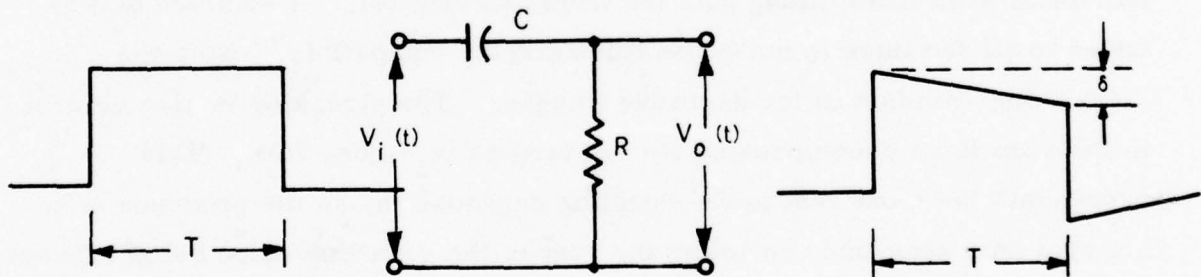
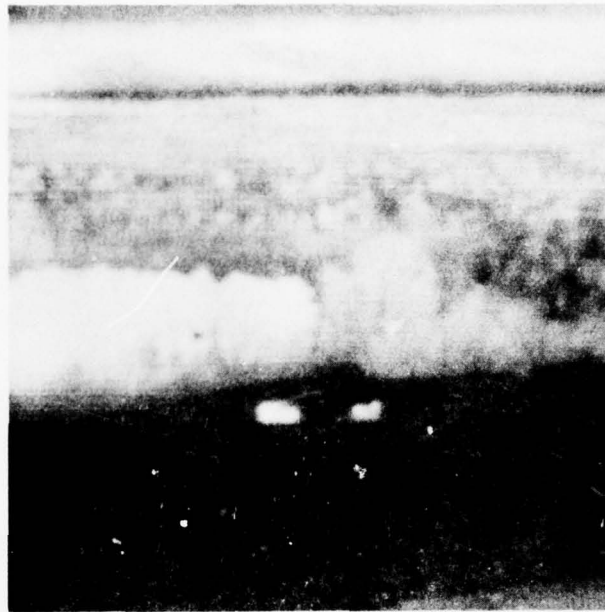


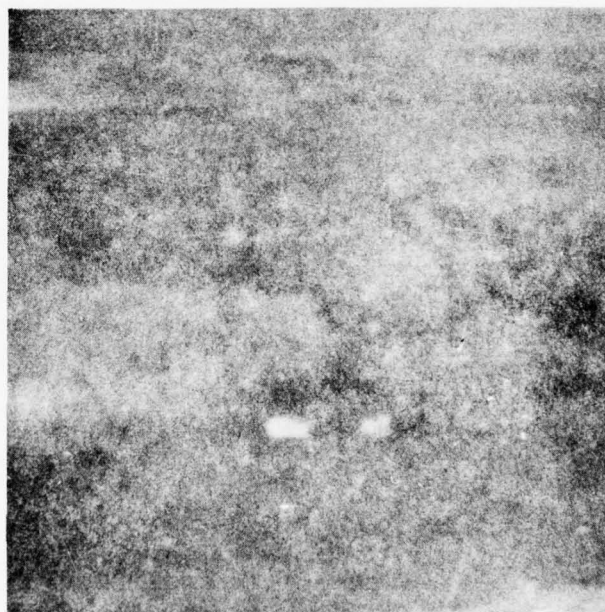
Figure 1. Rect Function Response of an RC Circuit

Much more serious than the transient effects, however, is that the parallel AC coupled detectors are not DC restored or clamped at the end of each scan line. When the image is viewing a stable scene, the average value of the video in each channel becomes zero independently of the neighboring channels. This has serious consequences when the average scene temperature is changing rapidly in a direction perpendicular to the scan because this difference is then lost in the display. A classical example is the loss of horizon definition because the detectors scanning the cold sky and those scanning the hot ground yield the same average video signal values. Figures 2 (a) and (b) show this phenomenon.

The second steady-state effect is the "streaking" produced by a very hot or cold target against a uniform background. This again happens in the parallel scan FLIR with no DC restore. Figure 3 shows this effect on test targets in Figure 3(a), degraded to approximate the loss of DC on the individual scan lines (along with the transient effects). A common bias is added to all the lines to make the video display compatible. Note the "streaking" evident in the degraded images. The streaking is also evident in the scan lines encompassing the hot targets in Figure 2(b). This represents the most severe AC coupling degradation, as the presence of a hot spot on a scan line can lower the rest of the scan line video below blacker than black on the display and obscure any detail present in the line. This happens because the average value of a scan line in the top of Figure 3(a) is lower than the average value of a line with the white-hot portion on it. With the loss of the DC, the average values of these lines in Figure 3(b) are now equal, depressing the lower line with respect to the upper one. This is also the reason for the shading on the hot and cold targets.

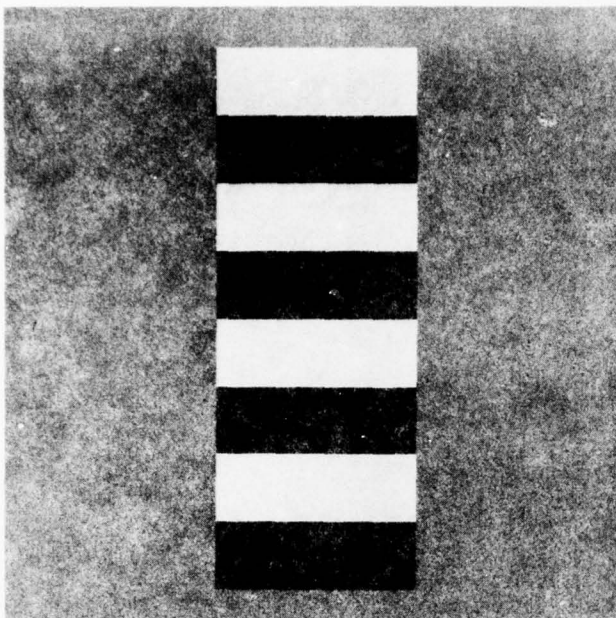


(a) Original

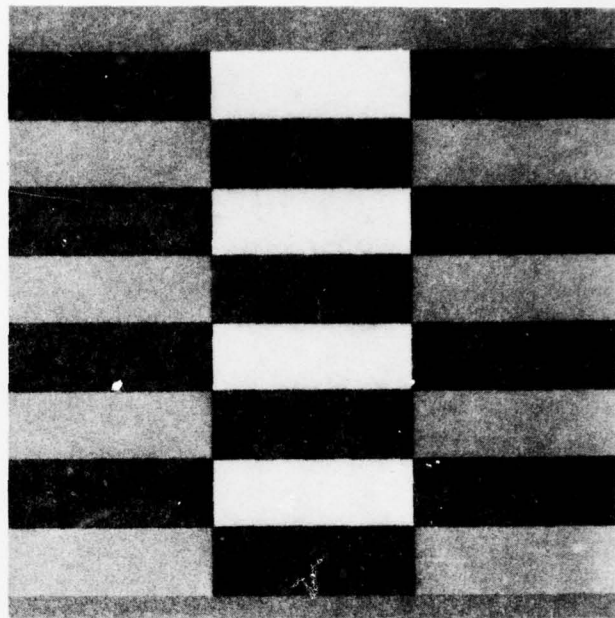


(b) Degraded

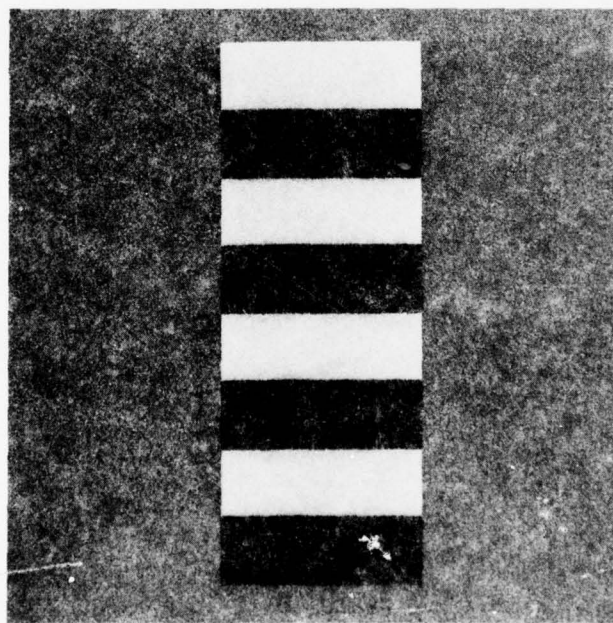
Figure 2. Effect of AC Coupling Degradation on a Typical DC-Restored FLIR Image



(a) Original test pattern



(b) Degraded due to AC coupling



(c) Synthetic DC restoration

Figure 3. Synthetic DC Restoration of Test Target 1

The most common DC restore technique, which corrects streaking, clamps the signals on all channels by imaging a common reference source prior to each scan. Each channel then measures scene radiance changes relative to the same reference level, and relative interchannel DC levels are preserved.

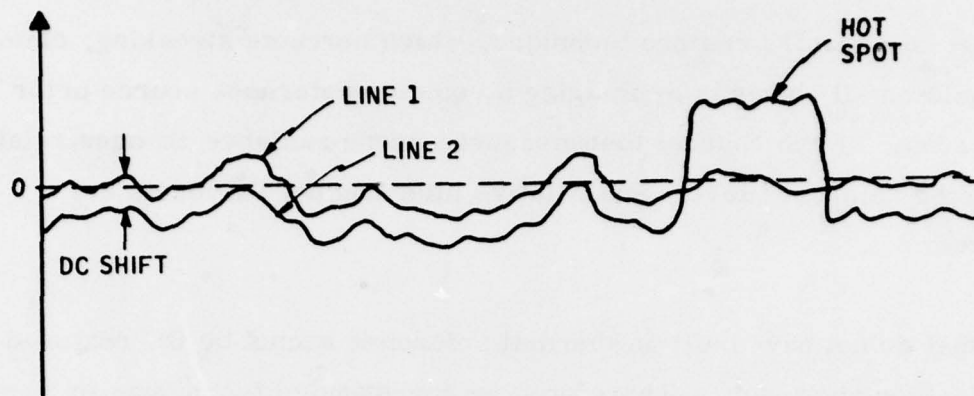
FLIRs that do not have built-in thermal reference cannot be DC restored in this fashion, however. Therefore, we investigated techniques to restore the line-to-line correlation after it has been lost by AC coupling. This uses the vertical line-to-line correlation present in most scenes.

A Synthetic DC Restore Algorithm *

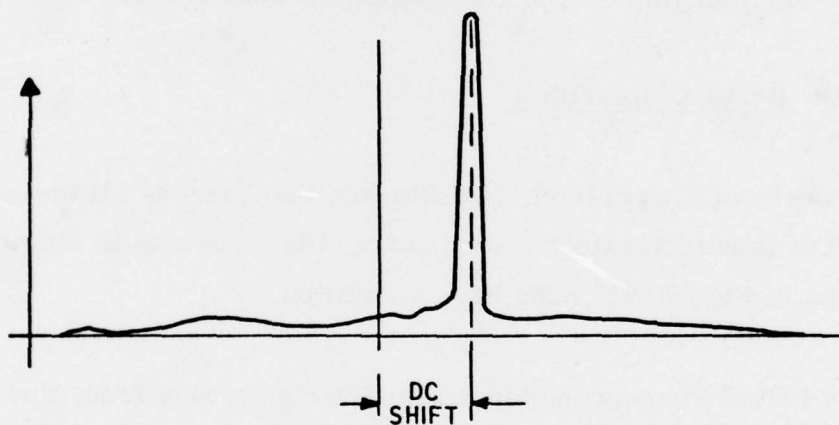
Although recovery of a completely lost horizon that extends all the way across the FOV is unfortunately a lost cause, the streaking in the presence of hot spots as in Figure 3(b) does have a solution.

We recognize that the background in a scene varies slowly from line to line. In the AC coupled video, the presence of a significant hot target in one line causes that line to be depressed with respect to the previous line. Therefore, the pixel-to-pixel differences of the two lines will be predominantly distributed at or near the average DC shift of the second line with respect to the first. Figures 4(a) and (b) illustrate this. By recognizing the peak in the histogram, we can identify the shift of the DC level and add it to the second line. The second line now serves as the reference for the next line, and so on down. Actually, we need only to detect the median of the histogram (see Figure 4(c)) to get an estimate of the DC shift. Figure 5 is a functional block diagram of this algorithm. A

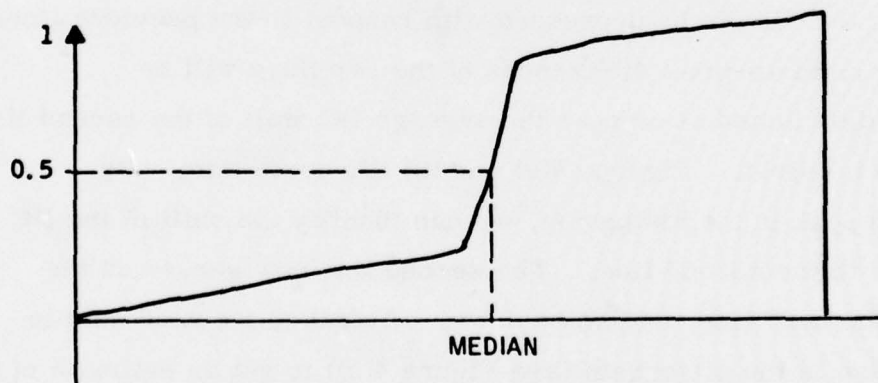
* "Pseudo DC Restoration Using Histogram Modification," MVL Patent Pending.



(a) Shows the Shifting of Background Level because of the Presence of a Very Hot Spot (Streaking)



(b) Histogram of the Difference of Line 1 and Line 2



(c) Cumulative Histogram of b

Figure 4. Line-to-Line Pixel Differences from AC Coupling

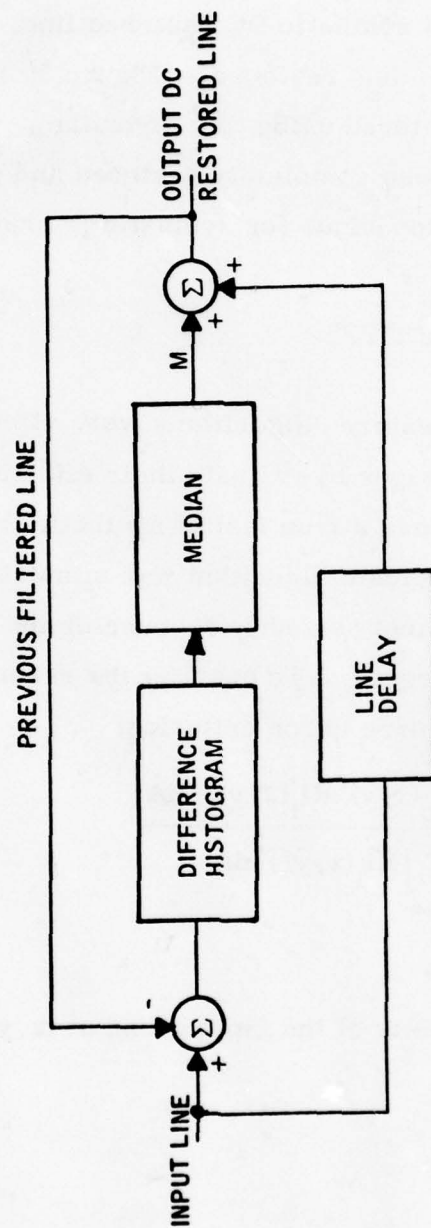


Figure 5. Functional Block Diagram of the Synthetic DC Restore Algorithm

difference histogram is formed by subtracting the current line from the previous filtered line. The median of the histogram is added to the current delayed line to get the current synthetic DC restored line. The line-to-line background correlation is thus restored. Figure 3(c) is the degraded test pattern of Figure 3(b) restored using this algorithm. We note that the line-to-line correlation has been completely restored and shading eliminated. This therefore is a powerful technique for synthetic DC restoration.

DC Restore Performance Evaluation

The candidate synthetic DC restore algorithms were simulated on degraded test targets and real FLIR images to evaluate their effectiveness. Figure 3(a) was chosen because it poses a true test of synthetic DC restore algorithms. In fact, the candidate algorithm was simulated on these patterns and succeeded in almost perfectly recovering the original line-to-line correlation of the test targets. To quantify the error in restoring we used the normalized mean-square error criterion:

$$NMSE = \frac{\iint [RI_o(x, y) - RI_I(x, y)]^2 dA}{\iint [RI_I(x, y)]^2 dA}$$

where

$RI_I(x, y)$ = Relative intensity of the input scene at (x, y)

$$= \frac{I_I(x, y) - I_I}{I_{IMAX} - I_{IMIN}}$$

and

$I_I(x, y)$ = Intensity at (x, y) of input scene

I_I = Average of $I_I(x, y)$ over the whole image

I_{IMIN} = Time average minimum of input image

I_{IMAX} = Time average maximum of input image

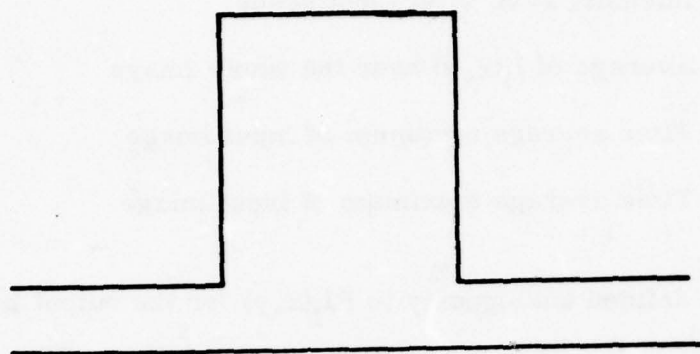
$RI_O(x, y)$ is defined analogously to $RI_I(x, y)$ for the output image.

The above criterion is insensitive to scaling and average DC level of the scene and will measure the true loss of line-to-line DC correlation. The performance of the algorithm as measured by the NMSE criterion was better than 1 percent for the image shown in Figure 3(c).

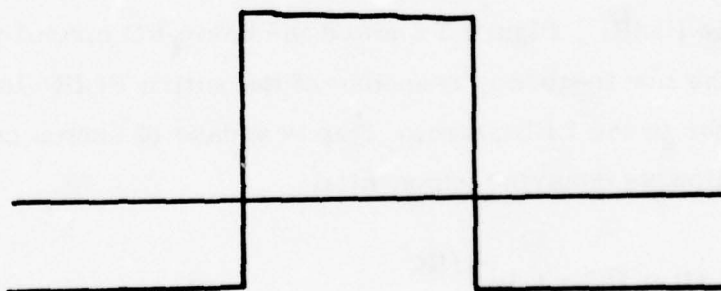
The above discussion pertained to the steady state effects of AC coupling in a parallel scan FLIR. We will now show that the transient phenomenon has a very minor effect on the video for the time constants encountered with the Common Module FLIR. Figure 1 showed the basic RC circuit that approximates the low frequency response of the entire FLIR electronics from the detector to the LED stage. Step response of such a circuit is given by the following decaying exponential:

$$U_O(t) = U(t - t_0)e^{-t/RC}$$

The response of the circuit to a Rect function (that more closely approximates a hot spot) can be derived from the unit step response and is shown in Figure 6. The important feature in this response is the droop δ of the top of the Rect function and the corresponding equal undershoot δ that follows



(a) Rect Function Input



(b) Degraded by Simulating the AC Coupling with
Common Module FLIR Parameters

Figure 6. Circuit Response to a Rect Function

the Rect function. The time constant of the equivalent RC circuit ($T = RC$) is related to the lower 3 dB cutoff frequency of the system (detector to LED) of approximately 8 Hz by

$$\tau = \frac{1}{f_{3dB}} = 1/8 \text{ sec}$$

The most serious degradation of this kind will obviously occur when the Rect function duration T is close to the scan line time, $1/60$ sec (for a parallel scan system). It is instructive to estimate the droop δ as a percentage of the Rect function for the Common Module FLIR:

$$\delta = (1 - e^{-T/\tau}) \times 100\%$$

where

$$T = 1/60 \text{ sec}$$

$$\tau = 1/8 \text{ sec}$$

$$\delta_{\max} = 12.5\%$$

which indicates that the transient DC droop and undershoot is not at all severe. This is because we very seldom would have the extreme case of a hot spot extending almost the entire line width.

Figure 6 shows the results of a simulation of the transient effect on a Rect function extending a third of a scan line. The computer simulation was done with recursive digital filters and the time constant was chosen to be $1/8$ sec, to model that of the Common Module FLIR. Note that the undershoot and

droop resulting from the transient are negligible. Consider, for example, a sharp hot object extending 10 percent of the scan line. This will have a transient undershoot δ following it of 1.3 percent of the hottest temperature of the target. This will not be observed on the display. In any event, this phenomenon will certainly not cause adjacent detail to be blacked out. Therefore, no transient correction is needed for the time constants encountered in this FLIR.

On the other hand, as we pointed out before, the steady state loss of DC on the individual scan lines is a much more severe effect. The synthetic DC restore algorithm that detects the DC level shift in adjacent scan lines is a very attractive one, as was shown both conceptually and through simulation.

It should now be apparent that synthetic DC restoration is needed with current generation parallel scan FLIRs without a built-in DC restore function (a temperature reference, for example). Serial scan FLIRs, on the other hand, do not have streaking in the presence of hot targets because there is no loss of DC from one scan line to the next as in parallel scan FLIRs. Therefore, the worst AC coupling effect these FLIRs can have is the transient variety we discussed above. Here it manifests itself as a gradual shading from the top to the bottom of the FOV. This is usually not very serious and can be easily corrected by the use of clamping.

Is DC Restoration Necessary in Second Generation FLIRs ?

With the advent of monolithic and hybrid focal plane detector/CCD processors, the previous reasons for AC coupling detector outputs may now be invalid. We will see below that DC coupling with background subtraction obviates all but one of the three reasons for AC coupling, that of the increased contrast sensitivity along a scan line. However, the two-dimensional contrast enhancement schemes developed and simulated in later sections do precisely that, without destroying the line-to-line correlation and degrading the vertical modulating transfer function (MTF), as would AC coupling of a parallel detector array. A detailed analysis of detector coupling in photovoltaic (PV) detector/CCD hybrid focal planes (of second generation FLIRs) was made in the interim report [3]. From consideration of dynamic range and fabrication complexity, DE coupling of PV detectors to the CCD was considered to be superior to AC coupling. Moreover, any advantage that AC coupling may have in increased dynamic range was easily offset by constant background charge subtraction with DC coupled focal planes. The last reason for the DC coupling practice, $1/f$ noise in detectors, was analyzed in the first interim report [1]. The conclusions were that $1/f$ noise is negligible in direct coupled PV/CCD detector interfaces. This was shown by considering a hybrid (PbSnTe) and HgCdTe) detector/CCD example. Even in the worst case of a fully parallel scan system (at 30 fps, 2:1 interlace), the $1/f$ noise energy (from 10^{-4} Hz to 250 Hz) is only about 4.5 percent of the total noise energy (10^{-4} Hz to 184 kHz) and hence would not be noticeable on the display.

In summary, direct coupling of PV detectors to the CCD is suitable for focal plane applications as long as the wavelength band and integration times

are compatible with reasonable CCD integration sizes. Excellent noise performance may be obtained. Background subtraction is of marginal utility if large ranges of background temperatures or large dynamic ranges are necessary, or if large variations in detector responsivity and/or saturation current exist. AC coupling is not practical for future FLIR systems requiring large numbers of detectors on the focal plane. $1/f$ noise is negligible for PV detectors, and two-dimensional contrast enhancement schemes can achieve increased contrast sensitivity without degrading the vertical MTF, as would AC coupling.

DETECTOR RESPONSIVITY EQUALIZATION

Uniformity of channel-to-channel responsivity in a multidetector system is an important requirement if high temperature resolution is to have any meaning. PV detectors can be fabricated at about 5 percent uniformity, but temperature resolution of order 0.05°K require uniformity of 0.1 percent or better among channels.

A detailed analysis of this problem for a parallel scan FLIR with direct coupled detectors was made in the interim reports [1, 3]. The nonuniformity of the detectors translates to differences in gain (to a ΔT) and offset (variation against a constant background) from detector to detector. Two approaches were suggested for responsivity equalization. One was a simple deterministic scheme involving a two temperature reference near the focal plane [3] for the gain and offset equalization. The necessary temperature source was realized by defocusing the references as imaged by the detector so that spot-to-spot temperature variation in the reference would not affect the calibration of the detectors. This relaxed uniformity of the two temperature sources was determined to be realizable with state-of-the-art technology.

An alternate (stochastic) approach to detector gain equalization was also developed [1]. This scheme was designed to be effective even when the temperature sources (above) were non-uniform. In effect, adjacent pairs of detectors were equalized with respect to each other in an iterative fashion. Because of the non-deterministic (stochastic) nature of the scheme, absolute look-to-look and spot-to-spot uniformity of the reference source is unnecessary. The scheme was simulated for a 15 element detector array assuming a detector gain variation with a standard deviation of 10 percent and the calibration source was assumed to have 5 percent one sigma variability from spot-to-spot and look-to-look. The iterative algorithm converged in 100 iterations (at 1/30 fps, this would mean 3 seconds). After convergence, the responsivities were equalized to within 0.1 percent. An extension of this idea for larger detector arrays was also developed.

In conclusion, channel responsivity equalization is a tractable problem. If uniform temperature references are designed into the field stop, for example, the process merely becomes a deterministic equalization. If the temperature source uniformity cannot be guaranteed, a stochastic approach will achieve the desired accuracy even with a variable reference source. Further work needs to be done, however, to determine whether the responsivity equalization can be done by dispensing with temperature references altogether. This would be invaluable in dithered sparse and staring array configurations where the uniform reference source would have to occupy the whole FOV. This would be amenable to a statistical approach. (Adjacent detectors see the same scene elements, in the long run, when the FLIR platform is in motion, and on the average should produce the same output.) This approach would use the first and second order statistics of the sensed video from each channel. The equalization scheme would be similar (but not identical) to the stochastic approach in the interim report [1].

SECTION III

STATISTICAL CHARACTERIZATION OF FLIR IMAGERY

In Sections IV through VII we will analyze the various image enhancement algorithms in light of their effectiveness for FLIR imagery. In order to quantify the effect of an enhancement process, however, we need measures of image quality that can be related to search effectiveness. Given this, we can compare algorithms by seeing how these quantifiers are transformed by each enhancement process. In measuring the image statistics we were also motivated by the need to analyze tactical FLIR imagery, with respect to shape, contrast and texture of typical targets and backgrounds. Such measures may provide useful discriminators in an automatic target screening system such as the Prototype Automatic Target Screener (PATs) that Honeywell is now building under NVL supervision.*

The statistics that were measured on each target and its surrounding background can be categorized as shape, intensity, and texture features. The detailed definitions of the various measures and how they were extracted can be found in the interim report [1]. For convenience, however, we will list the measures below and follow with a brief description of each measure.

* Contract #DAAK-70-77-C-0248

1. Shape Statistics (on target and target-like objects)
 - a. $\text{Perimeter}/\sqrt{\text{Area}}$
 - b. Number of edges
 - c. Histogram of the edge lengths (normalized by the perimeter (mean, standard deviation, skewness, excess))
 - d. Histogram of differential slopes of successive edges (mean, standard deviation, skewness, excess)
2. Intensity and Contrast Measures (on target and its background)
 - a. Intensity histogram of target (mean, standard deviation, skewness, and excess measured on the histograms)
 - b. Intensity histogram of its background (mean, standard deviation, skewness, and excess measured on the histograms)
 - c. Average contrast of target/background
 - d. Peak contrast of target/background
 - e. Histogram of Sobel gradient operator on the edges
 - f. Histogram of the gradient across the edge--edge contrast measures
3. Texture Features

These are grey-level difference histograms in four principal directions, and measures computed from these histograms. Texture features are measured on target and background.

These features were measured on the original and enhanced FLIR imagery. A repertoire of analysis and display tools were developed to analyze the statistics. They are histogram routines to generate and display the distribution of the statistics and compute various moments, and two-dimensional scatter plots to display the cluster plots of the features taken one pair at a time. The scatter plot programs are interactive and versatile enough to plot and display any pair of up to 80 features on an ensemble of 120 targets. Annotated symbols delineate the target and background classes.

We will now briefly describe each feature in turn and illustrate with plots of the statistics taken from 90 targets in thermal images supplied by NVL.

SHAPE STATISTICS

All shape statistics are derived from an object-boundary tracing routine and the resulting chain code which describes the perimeter of the object. The boundary extraction routine uses the Sobel gradient operator. Details can be found in the first interim report [1].

Figure 7 shows a line printer plot of a tank extracted by this technique. Pixel intensities interior to the target are negated for ease of separating target from background when extracting brightness and texture features. The notation (10 x 20/137) gives the (height x width/area) of the target in pixels.

Histograms of these measures (such as edge lengths and differential slopes) were computed for each target. Since these histograms are unwieldy for comparing across an ensemble of targets, they were further reduced to their moments. Specifically:

Mean	$\mu = E(x)$
Standard deviation	$\sigma = E(x - \mu)^2$
Skewness	$\frac{E(x - \mu)^3}{\sigma^3}$
Excess	$\frac{E(x - \mu)^4}{\sigma^4} - 3\sigma^4$

Two-dimensional scatter plots of these target shape features across 45 original (NVL supplied) thermal images are shown in Figures 8, 9 and 10. The classes of targets were T (tank), A (Armored Personnel Carrier - APC), and J (jeep), and we see that there is considerable overlap among the target classes in all of the shape feature plots, indicated that these shape features alone are not complete discriminators as to target class. There was also a preponderance of tanks in the data set that biases any statistical conclusion we derive from the ensemble. Note, however, that tanks have lower mean slopes (which means larger number of shorter sides than the other targets) in Figure 9. Figure 10 substantiates this in the mean edge length feature.

Table 1 summarizes these shape features (means and standard deviations) across the ensemble of all 45 images from the same sensor (NVL thermoscope) and for each class of targets. From this, the P/\sqrt{A} feature appears to be useless as a discriminator. The mean number of edges shows a slight separation in the means but the standard deviation within each

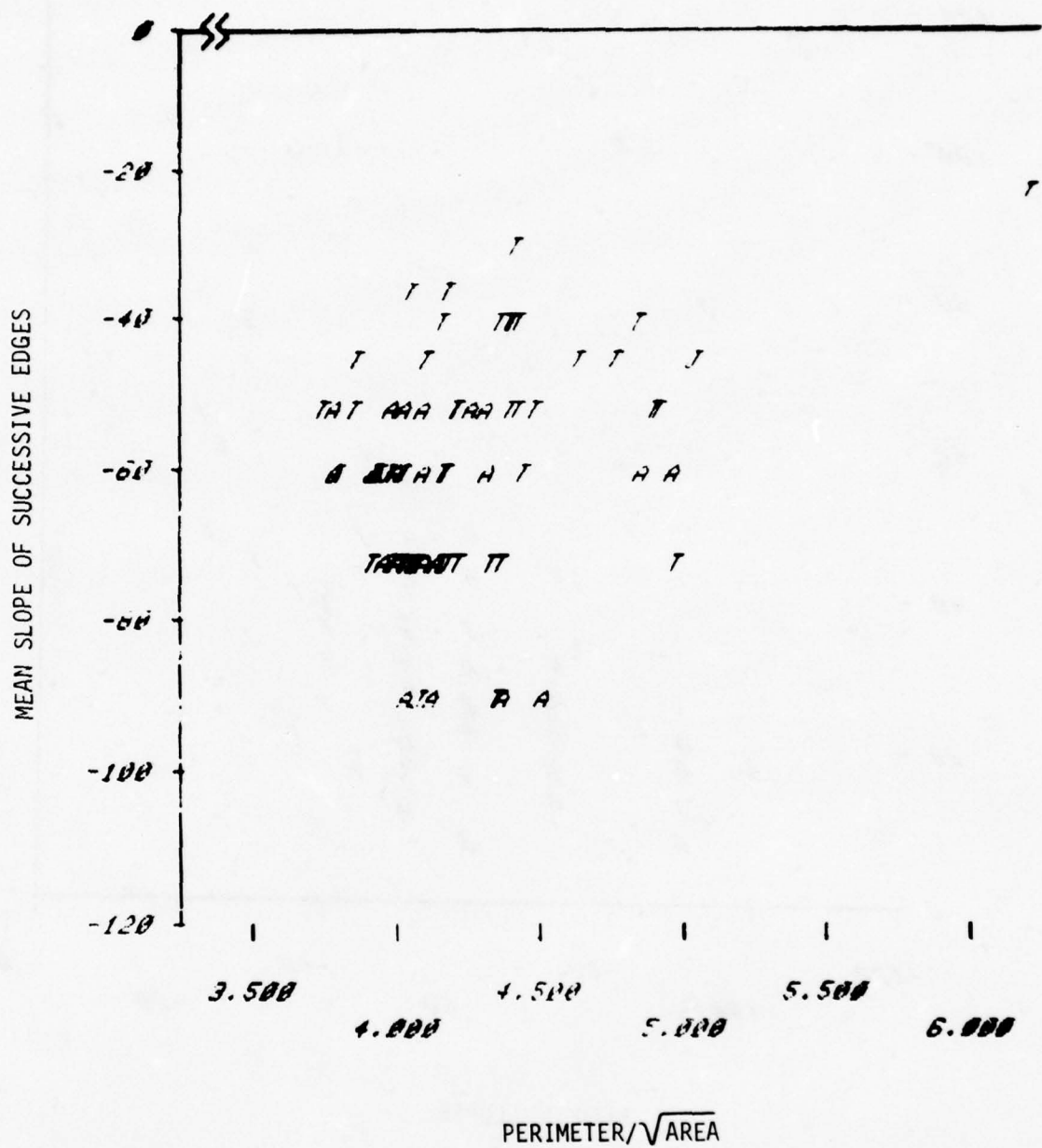


Figure 8. Mean Slope of Edges vs. $\text{Perimeter}/\sqrt{\text{Area}}$
 T = tank; A = APC; and J = jeep

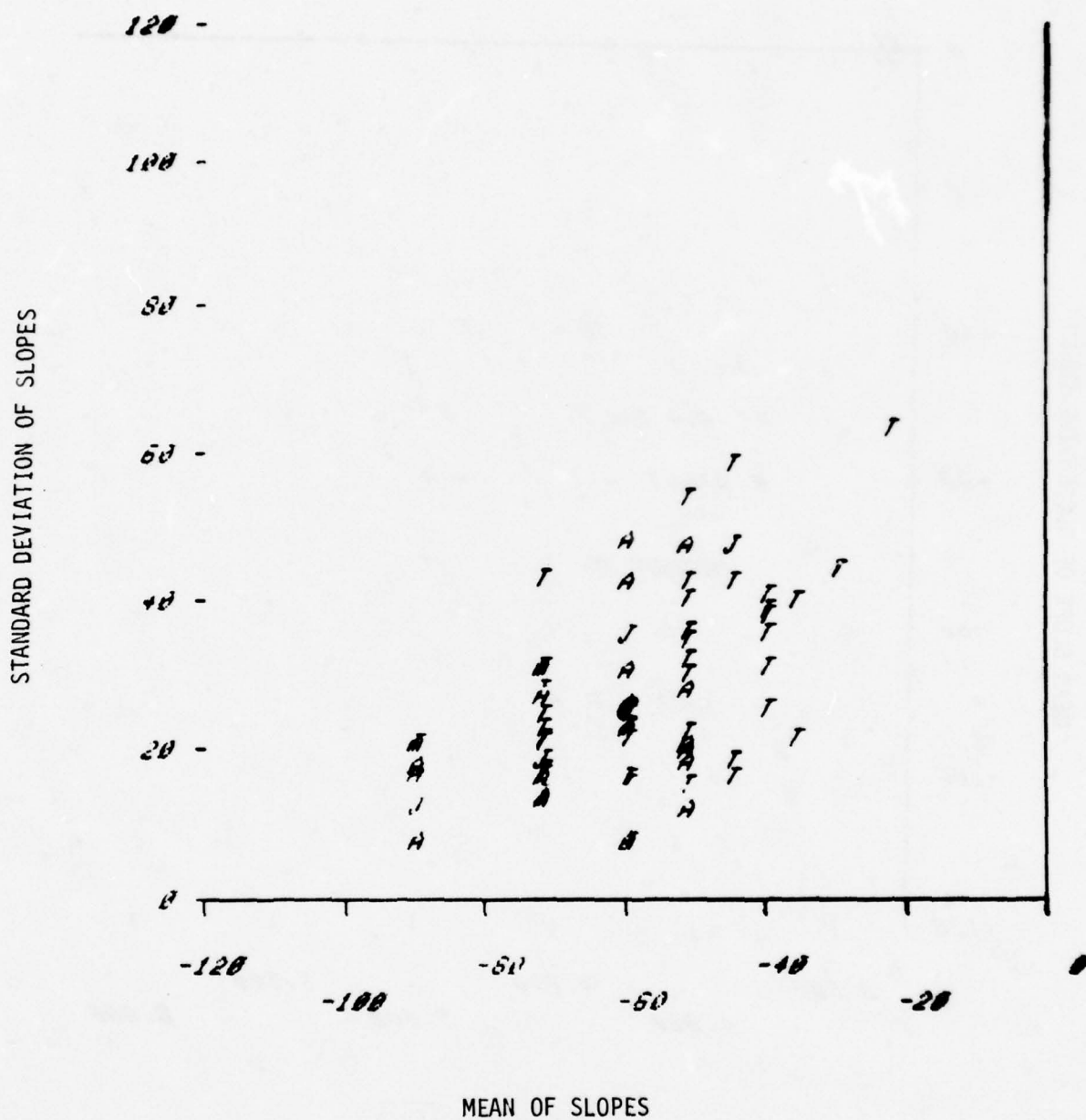


Figure 9. Mean vs. Standard Deviation of Relative Slopes

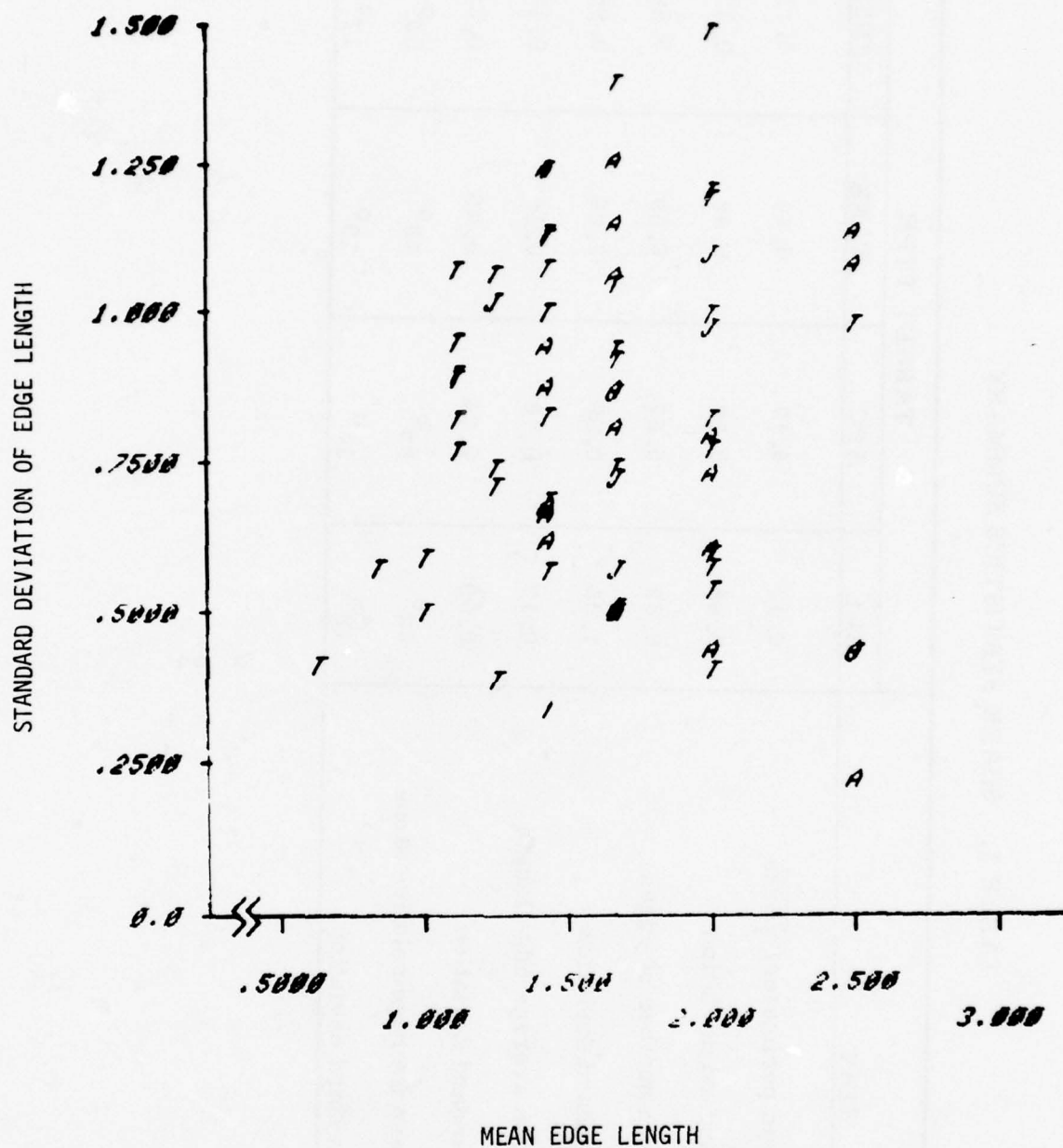


TABLE 1. SHAPE STATISTICS SUMMARY

STAT	TARGET TYPE			
	ALL	APC	TANK	JEEP
Mean perimeter/ $\sqrt{\text{Area}}$	4.32	4.17	4.40	4.11
Standard deviation	0.43	0.28	0.46	0.32
Mean number of edges	6.32	5.67	6.89	5.60
Standard deviation	1.95	0.99	2.29	0.96
Mean average edge length	0.17	0.18	0.16	0.18
Standard deviation	0.05	0.04	0.05	0.03
Mean average relative slope	62°	66°	58°	65°
Standard deviation	18°	13°	19°	11°

class exceeds the between-class separation. The average edge length and average differential slope features appear to discriminate between tanks and APCs but offer little against jeeps. (Since there were only four jeeps in the set, this is of questionable significance.) This shows the utility of scatter plots such as Figures 8, 9 and 10 over the ensemble summaries such as Table 1, which do not always tell the whole story.

INTENSITY AND CONTRAST STATISTICS

Once the target boundary has been found by the boundary tracing algorithm, a box containing the target is extracted from the image, as in Figure 7. Intensity and texture statistics are measured on the target (within the boundary) and the background surrounding it and contained in the box. To ensure that the background is adequately represented, the box surrounding the target allows a liberal margin (at least 20 pixels) all around the target.

Histograms of this target and background intensities were measured on every target box. As with the shape features, these histograms are further reduced to the four moment functions derived from the histograms, i.e., mean, standard deviation, skewness, and excess. Figures 11 and 12 show typical intensity histograms of the target and its background. Figure 13 shows mean versus standard deviation of target (T, A, J) and background (B) intensities for all 90 targets in the set. We see that no single threshold in the mean or standard deviation of the intensity will separate all of the targets from their backgrounds. This is to be expected, of course, as the backgrounds can encompass almost as large a dynamic range as targets over a large ensemble of images. Figure 14 reveals an interesting fact. The target intensities are predominantly negatively skewed, while their

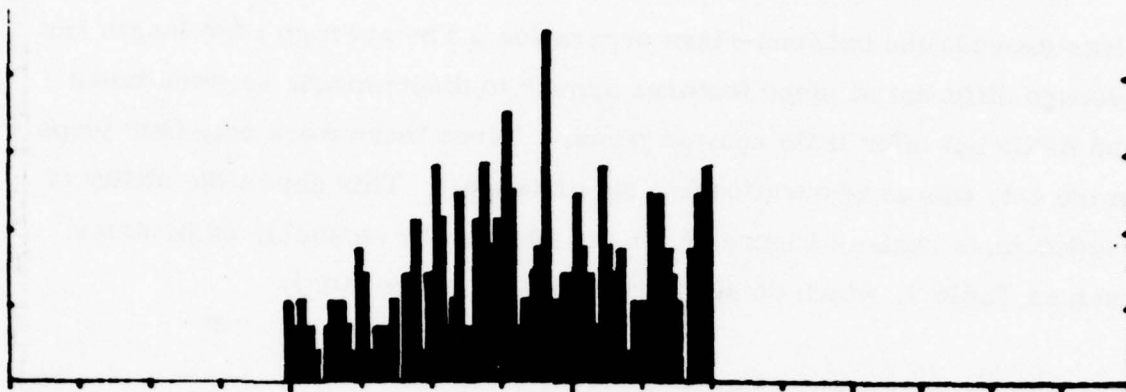


Figure 11. Target Intensity Histogram

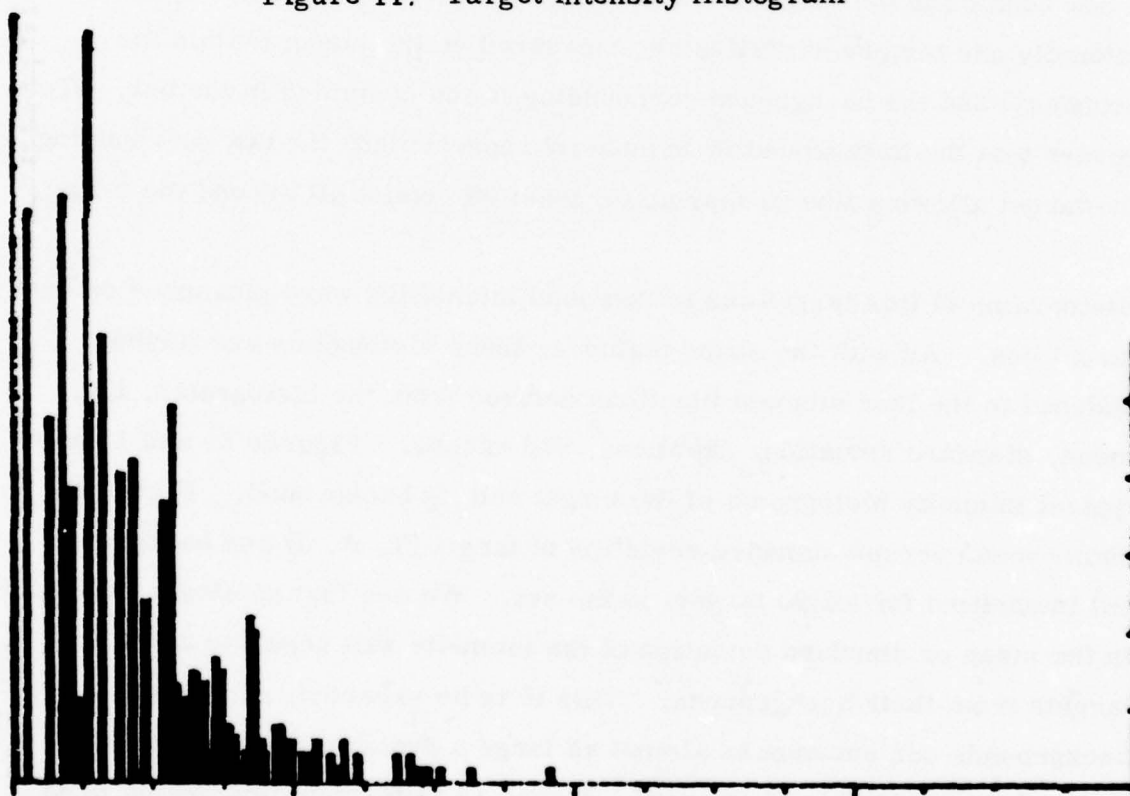


Figure 12. Background Intensity Histogram

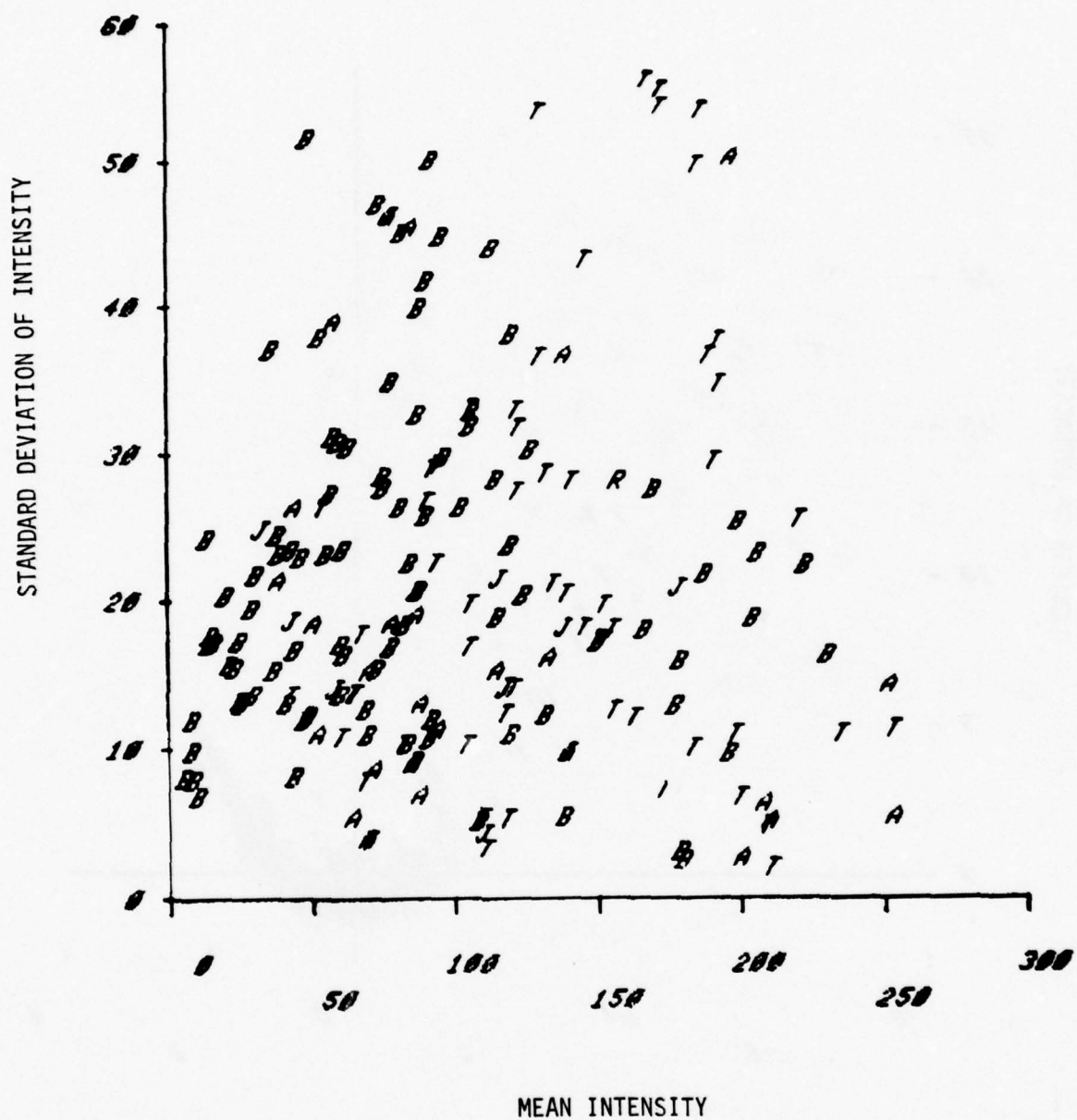


Figure 13. Mean and Standard Deviation of Intensities of Targets (T) and Backgrounds (B)

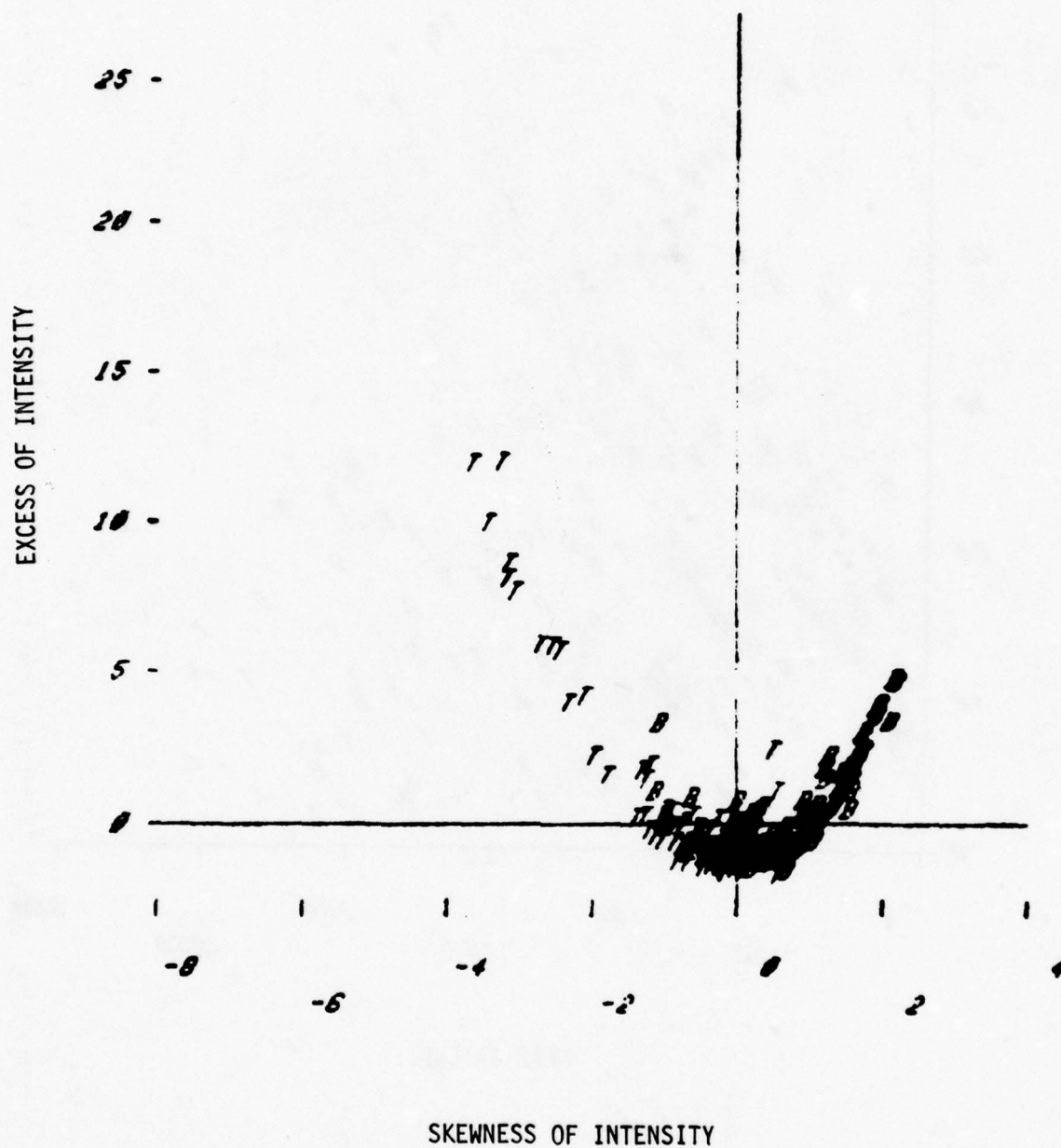


Figure 14. Skewness vs. Excess of Target (T) and Background (B) Intensities

backgrounds exhibit a positive skewness in almost all cases (see Figures 11 and 12, for example). In fact, of the 90 targets, less than nine targets have positive skewness and nine backgrounds have negative skewness. This is easier to see in Figure 15 which is a plot of the intensity mean versus skewness. Although not separable in the mean intensity, the targets and backgrounds exhibit remarkable separability in the skewness dimension. Note, however, that this property cannot be used per se for automatic scene segmentation because the target and backgrounds intensity statistics here assume that the targets have been already isolated.

These intensity histograms (target and background) also yield a class of contrast statistics. These are the average and peak contrast measures, defined below:

$$\text{Average contrast} = \frac{|M_T - M_B|}{M_T} \quad \text{or} \quad \frac{|M_T - M_B|}{M_B}$$

$$\text{Peak contrast} = \frac{|I_{T, \max} - M_B|}{I_{T, \max}} \quad \text{or} \quad \frac{|I_{T, \max} - M_B|}{M_B}$$

where

T refers to the target, B to the background, M is a mean and $I_{T, \max}$ is the peak target intensity.

The normalization can be with respect to either target or background, as appropriate. Both definitions were coded and the corresponding statistics gathered.

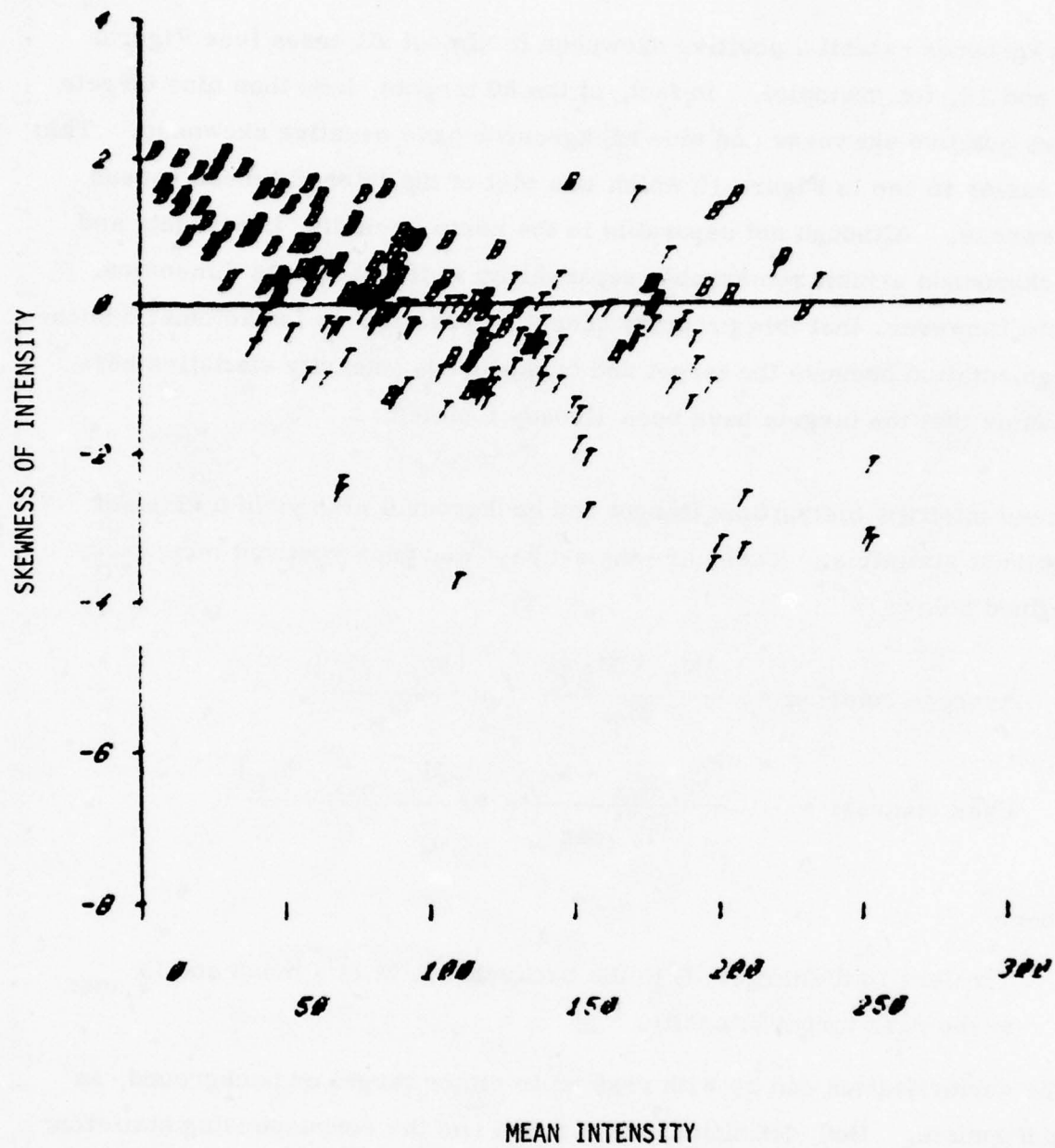


Figure 15. Mean vs. Skewness of Target (T) and Background (B) Intensities

In addition to these measures, an edge contrast measure was also computed. This is the average of difference of intensities of pixels immediately adjacent to the boundary on either side of the boundary. This is normalized with respect to the mean target intensity and is an indicator of the edge sharpness. The main driver behind these contrast statistics is that they are a measure of target conspicuity and are related to several search effectiveness models. Image enhancement algorithms--contrast enhancement algorithms in particular--can be judged against one another by the effect they have on these measures, as we will see in subsequent sections.

Table 2 summarizes some of the intensity and contrast statistics gathered on the set of 90 thermoscope images supplied by NVL. As seen in Table 2, we also measured the statistics on subsets of these images that had similar characteristics. The notations NVL1, NVL2, NVL3, NVL4, NVL5 refer to these subsets. NVL1 and NVL2 comprise 40 daytime thermoscope images (tanks, APCs, and jeeps). NVL3 (11 images) primarily contains close range tanks and APCs; NVL4 and NVL5 are mainly far range low contrast (some times with cold targets) thermoscope images of tanks (and an occasional cow). We see in Table 2 that there is a wide variation in the target contrasts over the set of images (as measured by the standard deviation). These contrast features are used in later sections to quantify the effects of the various image enhancement functions.

TABLE 2. SUMMARY OF CONTRAST STATISTICS
BEFORE ENHANCEMENT

STATISTICS	ALL CLASSES	NVL1 NVL2	NVL3
1. Mean of average target intensities	126	110	191
2. Standard deviation of average target intensities	53.7	43.7	48.1
3. Mean of average background intensities	84.0	57.3	1.58
4. Standard deviation of average background intensities	55.5	31.8	49.0
5. Mean of average contrast with respect to target	0.425	0.498	0.181
6. Standard deviation of average contrast with respect to target	0.239	0.205	0.109
7. Mean of average contrast with respect to background	1.19	1.55	0.26
8. Standard deviation of average contrast with respect to background	1.50	1.67	0.23
9. Mean of peak contrast with respect to target	0.497	0.604	0.222
10. Standard deviation of peak contrast with respect to target	0.252	0.204	0.140
11. Mean of peak contrast with respect to background	2.32	3.03	0.37
12. Standard deviation of peak contrast with respect to background	3.33	3.76	0.40
13. Mean of edge contrast	0.249	0.343	0.101
14. Standard deviation of edge contrast	0.233	0.168	0.072

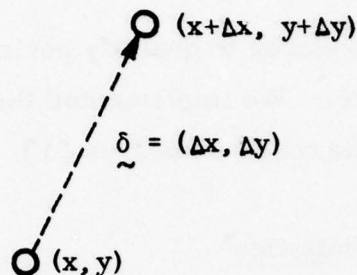
TEXTURE FEATURES

Texture features are measures to quantify periodic grey-level variations in the image, called texture. We implemented the grey-level difference statistics based on the Haralick measures [1].

Grey-Level Difference Statistics

Assume that the texture is to be measured over a local area of the image. Consider all pairs of points in the region, exactly at a vector distance $\delta = (\Delta x, \Delta y)$ apart (as in Figure 16). Let $|p(x+\Delta x, y+\Delta y) - p(x, y)|$ be the grey-level difference for each such pair. The histogram of these grey-level differences (of all pairs of points exactly δ apart) is called the grey-level difference histogram $p_\delta(\cdot)$. If there are N grey levels, then there are N bins in this difference histogram. This is a measure of the probability density of grey-level differences occurring in the image at a given spacing and orientation. The shape of this histogram is a measure of the texture. The various descriptors (describing this histogram) are:

Mean	$\frac{1}{N} \sum_i i p_\delta(i)$
Contrast	$\sum_{i=0}^{N-1} i^2 p_\delta(i)$
Angular Second Moment	$\sum_{i=0}^{N-1} p_\delta^2(i)$
Entropy	$-\sum_{i=0}^{N-1} p_\delta(i) \ln[p_\delta(i)]$



$$\begin{aligned} \tilde{\delta} = \{ & [0, 1), 1, 0), (1, 1), (1, -1)\} \\ & \{ [0, 2), (2, 0), (2, 2), (2, -2)\} \dots \end{aligned}$$

Figure 16. Illustrating Point Pairs for Grey-Level Difference Statistics

Figure 17 gives the grey-level difference histograms for a target and its immediate background area for $\tilde{\delta} = (0, 1), (2, 2), (4, 0)$ and $(8, 8)$. Note that $\tilde{\delta} = (2, 2)$ and $(4, 0)$ show the greater differences between the target and the background than $\tilde{\delta} = (0, 1)$ or $\tilde{\delta} = (8, 8)$.

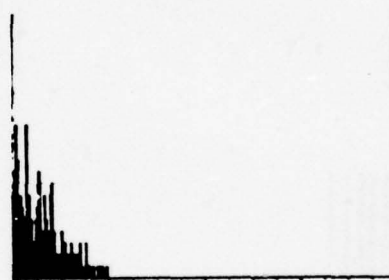
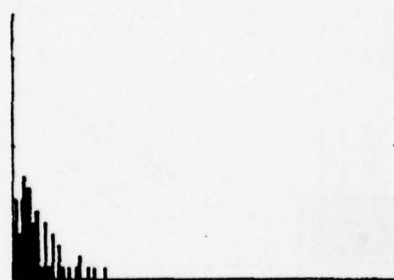
The difference histograms offer a great deal of information on the texture. But we need fewer descriptors of texture than the full histograms. Hence the above measures (mean, contrast, etc.) were computed from these histograms. These were computed on the target (tanks) and background areas for the NVL supplied thermal images; i.e., for $\tilde{\delta} = [(0, \Delta), (\Delta, 0), (\Delta, \Delta), (\Delta, -\Delta)]$ corresponding to orientations of $0^\circ, 90^\circ, 45^\circ$, and -45° and spacings of Δ .

TARGET

BACKGROUND



$\delta = (0, 1)$

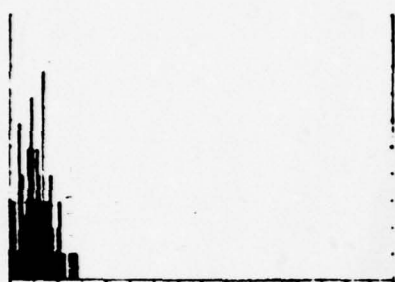


$\delta = (2, 2)$

Figure 17. Grey-Level Difference Histograms for a Representative Target and its Background for Various Values of δ

TARGET

BACKGROUND



$\delta = (4, 0)$



$\delta = (8, 8)$

Figure 17. Grey-Level Difference Histograms for a Representative Target and its Background for Various Values of δ (Concluded)

Interpretation of the Texture Features

The grey-level differences at a given direction and spacing are a measure of the periodic activity in the image at that spacing. If the texture were coarse, for example, we would see a larger value in the "mean" texture feature for the larger spacings.

Texture for small spacings ($\Delta=1$) also has significance in estimating noise in the image, however. To see this, consider a slowly varying background region in the image. The contrast feature for $\Delta=1$ computed over this region is a good measure of the noise variance in the image. In fact, under the assumption that the scene variation between adjacent pixels is small and the noise is white, we can show that the noise variance σ^2 on the image is given by

$$\sigma^2 = 1/2 \text{ CONTRAST } (\Delta=1)$$

This is a much better estimate of the image noise variance than if we computed the variance of the image background assuming the scene background to be constant. This is because the texture measure allows the scene to vary slowly over the background, whereas the image variance will include the true background scene variance in the estimate of the image noise. The above definition was used to measure image noise subsequently in quantifying signal-to-noise ratio in the FLIR and thermoscope images we analyzed (see Section 8).

The above texture measures--primarily mean and contrast measures (the first and second moments of the grey-level difference histogram)--were measured on the 90 targets and their backgrounds. Figure 18 shows the plot of the two features ($\Delta = 4$) over all the backgrounds (B) and targets (A, T, J, R). We see that they are highly correlated according to the square law. This is as it should be because the mean feature is the expected value of the absolute difference, whereas the Contrast feature is the expected value of the differences squared. Since these features are highly correlated we show only the mean feature in all subsequent texture examples. Figures 19 through 22 are two-dimensional plots of the mean texture feature versus the mean intensity of backgrounds and targets for spacings of $\Delta = 1, 2, 4$, and 8, respectively. Looking at the texture feature axis only as we go from $\Delta = 1$ to $\Delta = 4$, we see progressively greater separation between backgrounds and targets. The separation for $\Delta = 8$ is once again poor, indicating that maximum separability in texture is obtained around $\Delta = 4$. This is substantiated by the fact that for small Δ ($\Delta = 1$), the texture feature is reflecting mere noise. Very large Δ ($\Delta = 8$) on the other hand, is already of the order size of the target and therefore does not have much discriminatory value. The shape of the grey-level difference histograms in Figure 17 also shows that the histograms differ the most in shape from background to target for $\Delta = 4$.

In summary, we defined various measures to quantify target shape, contrast with respect to target, and texture, in targets and their backgrounds. These were measured on a set of 90 FLIR targets and their backgrounds and examples of their statistical plots were given. The shape descriptors were found to be unreliable for discriminating between the targets classes. The reason for this was the extremely large within-class variance exhibited

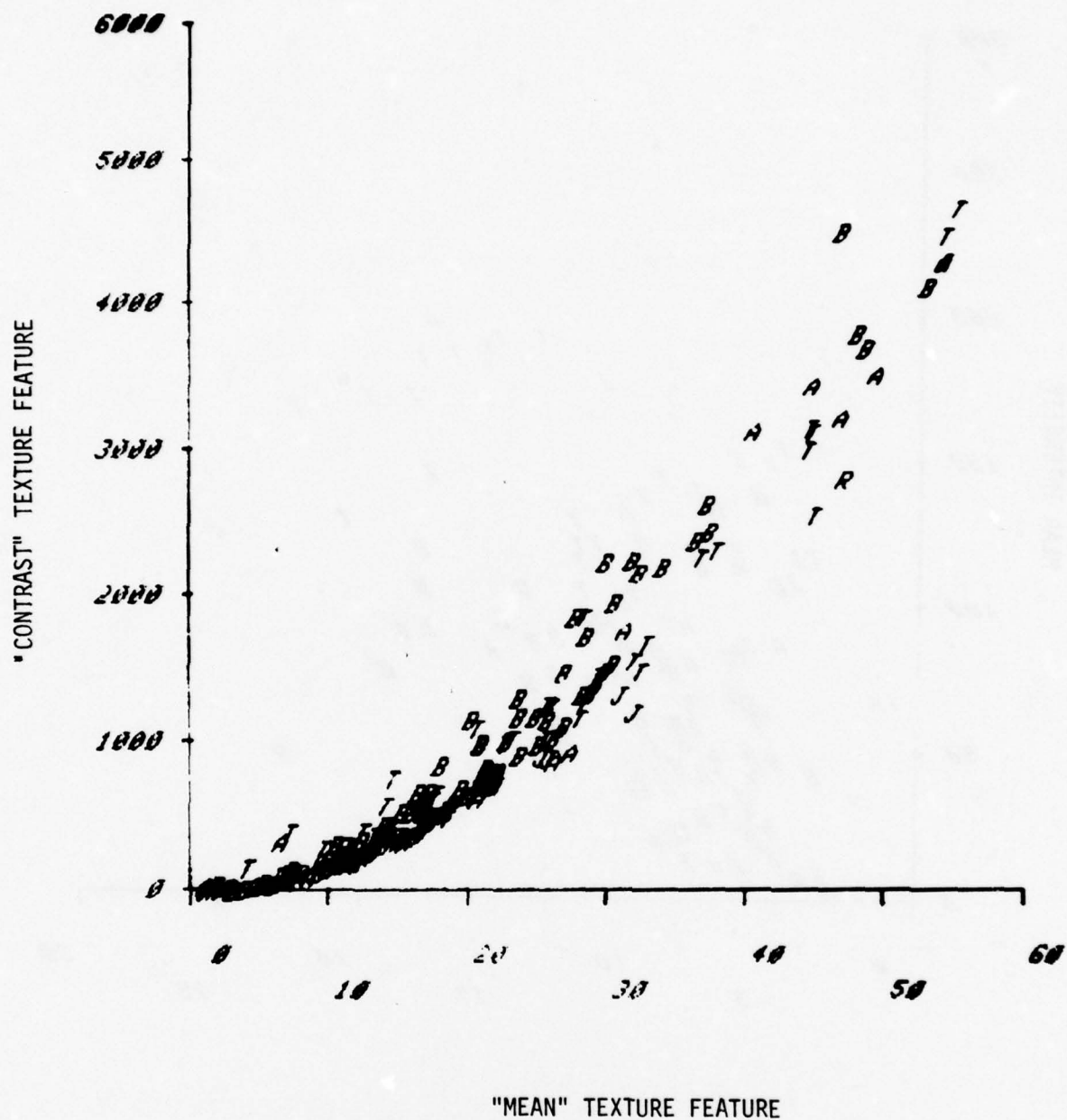


Figure 18. "Mean" vs. "Contrast" Texture Features for Spacing $\Delta = 4$

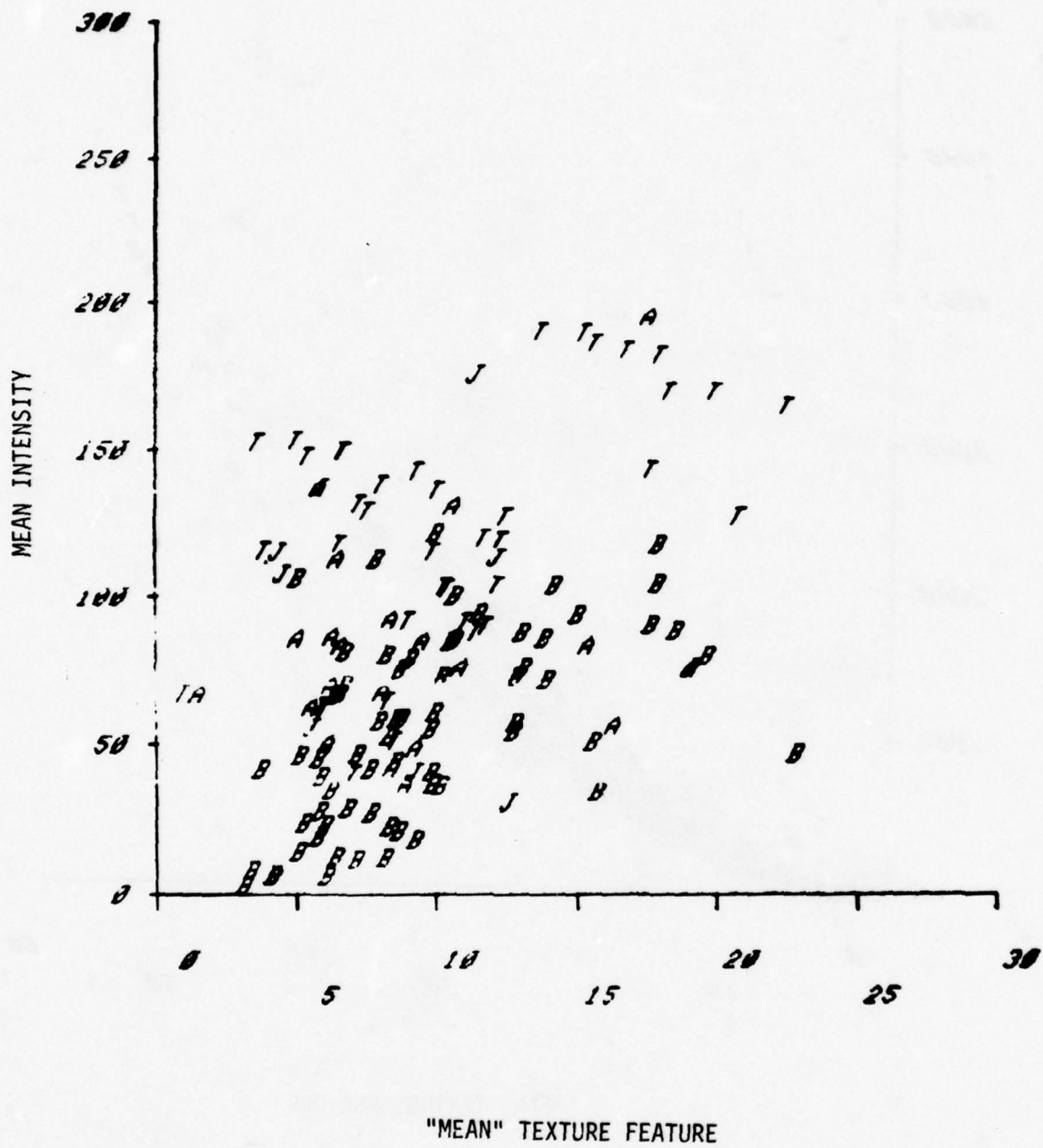


Figure 19. Mean Intensity vs. "Mean" Texture Feature for Spacing $\Delta=1$, Background (B), Targets (T, A, J)

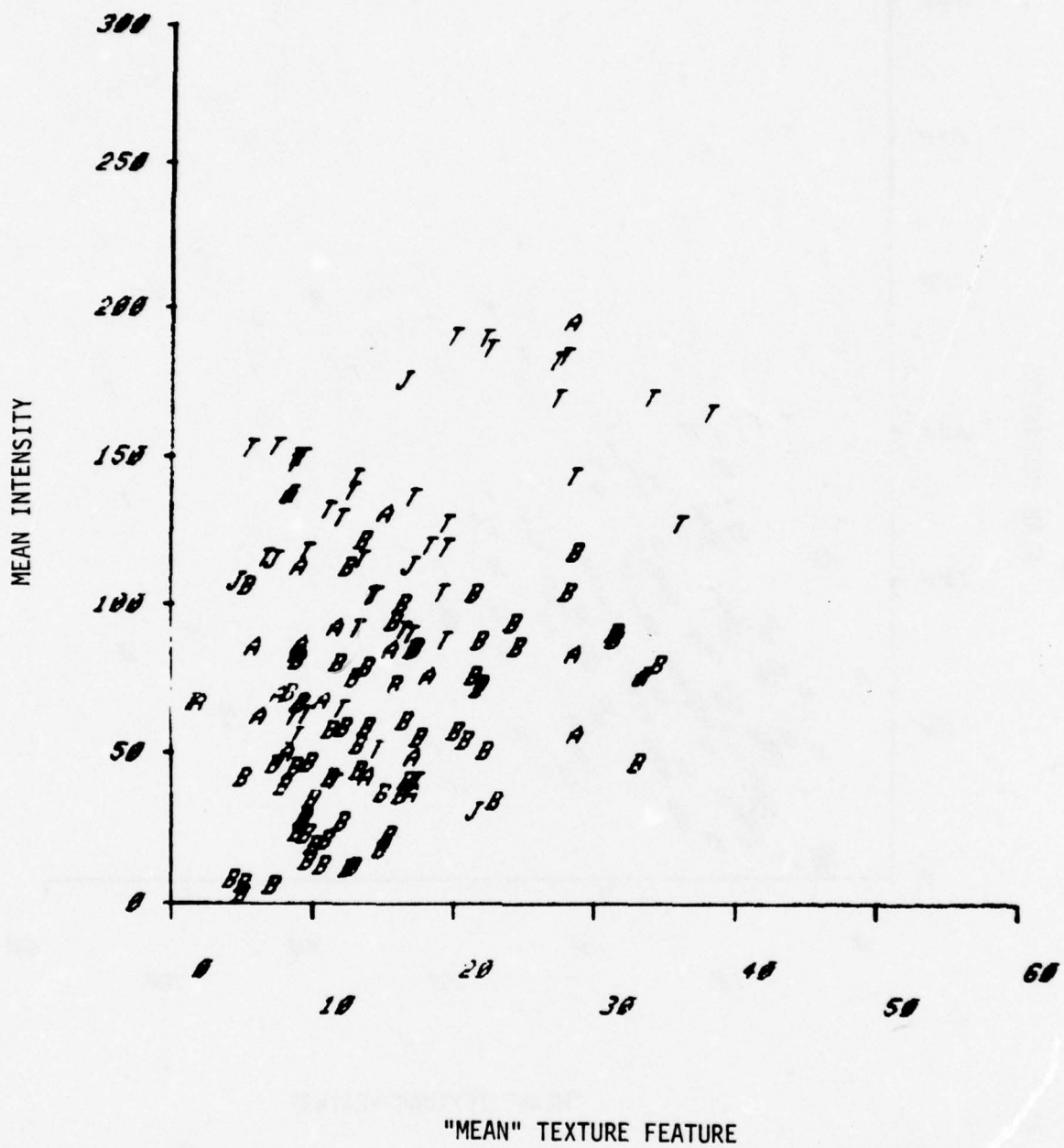


Figure 20. Mean Intensity vs. "Mean" Texture Feature for Spacing $\Delta=2$

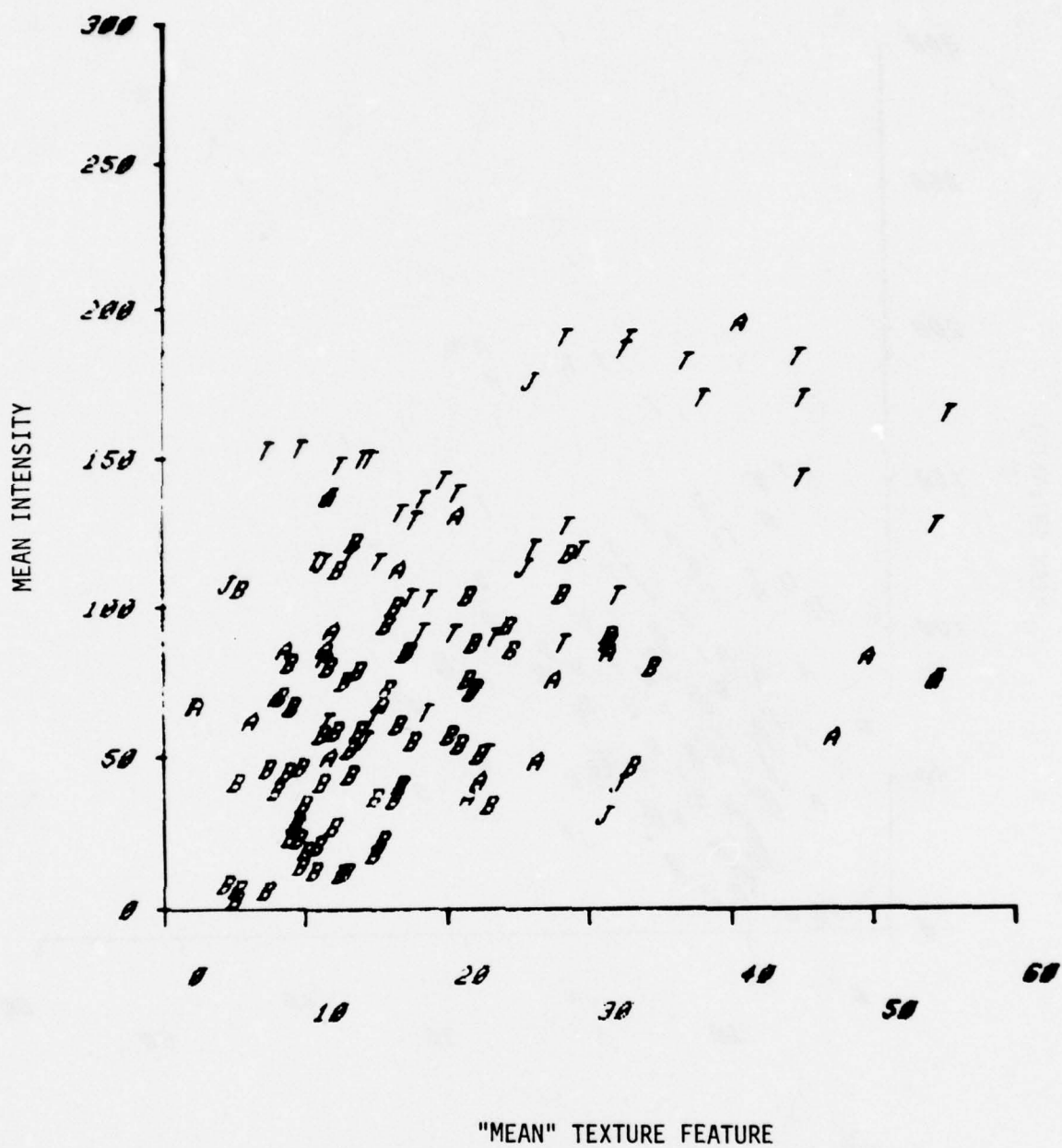


Figure 21. Mean Intensity vs. "Mean" Texture Feature for Spacing $\Delta = 4$

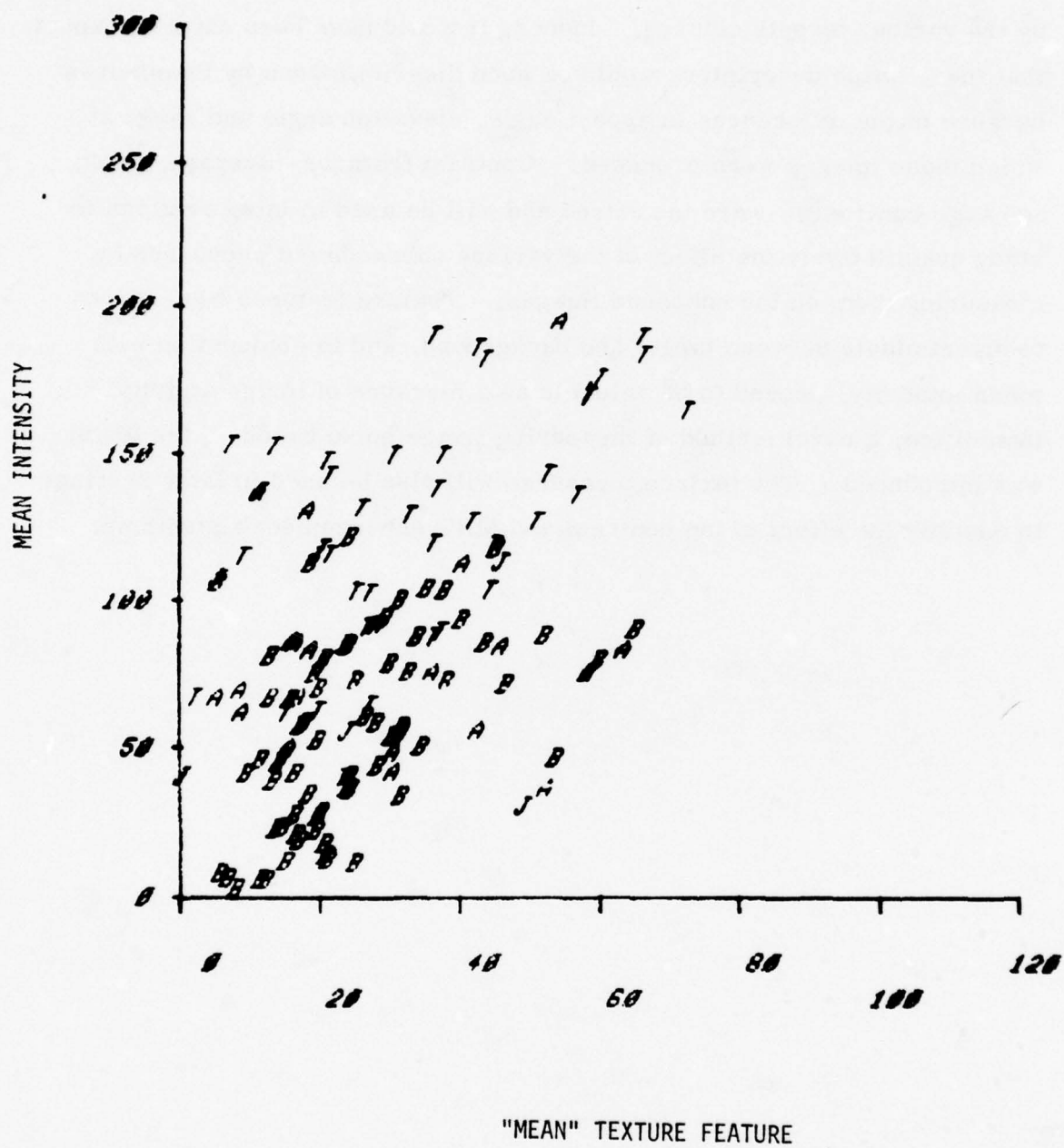


Figure 22. Mean Intensity vs. "Mean" Texture Feature for Spacing $\Delta=8$

by the various targets classes. Indeed, it would have been naive to expect that these shape descriptors would be good discriminators by themselves because of the differences in aspect angle, elevation angle and range at which these images were produced. Contrast features--average, peak, and edge contrasts--were measured and will be used in later sections to study quantitatively the effect of the various enhancement processes by measuring them on the enhanced images. Texture features were shown to discriminate between target and background, and in conjunction with mean intensity, proved to be valuable as a measure of image activity. In addition, a novel method of measuring image noise based on the texture was introduced. The texture measures will also be used in later sections to quantify the effect of the contrast and MRT enhancement algorithms.

SECTION IV

CONTRAST ENHANCEMENT

INTRODUCTION

This section reviews the contrast enhancement objectives, outlines the algorithms investigated, and discusses their effectiveness with the aid of the imagery statistics defined in Section III. Several contrast enhanced thermal images illustrate the algorithms.

THE NEED FOR CONTRAST ENHANCEMENT

Thermal imagers have a high inherent scene dynamic range. For example, the dynamic range of a FLIR scene with a cold sky and hot ground background can be as high as 1000:1, with a 100° temperature range and 0.1° MRT. With the advent of higher-sensitivity detectors the dynamic range could be even greater. This high dynamic range video information must be displayed on an imager display that has a luminance range typically of 15:1 to 30:1. If saturation in some part of the display and blackout in others is to be avoided, the gain of the display should be set very low. This means that low contrast (temperature difference) information in the scheme is translated to even smaller luminance (brightness) differences on the display, and cannot be perceived by the operator as it falls below the contrast sensitivity threshold.

It is possible, of course, to adjust the gain and brightness controls on a FLIR imager to selectively expand a given temperature range of interest but this involves extensive interactive manipulation of the controls and is therefore not desirable. Contrast enhancement algorithms were investigated in order to enhance the local contrast of the thermal images without exceeding the display dynamic range in a totally "hands-off" mode. Emphasis was on algorithms that would be capable of being implemented in real time in second generation FLIRs. After an extensive literature survey, a number of promising algorithms were identified, simulated, and evaluated. They were:

1. Automatic global and local area gain and brightness control (LAGBC).
2. High frequency emphasis spatial filters. (Edge enhancement, crispening and homomorphic filters).
3. Point transforms for grey-scale mapping (Global and local area histogram modification).

The main thrust of these techniques was to enhance the local (such as within-target and target/background) contrast in the thermal scene while compressing the overall dynamic range of the scene to within the display luminance limits. These algorithms were coded and tuned for optimum performance on FLIR imagery. Of these algorithms, three were selected to enhance a set of over 40 thermal images. The imagery statistics gathered over the enhanced images were then used to quantify their performance. These enhanced images were also input to the human factors enhancement evaluation task (Section VIII).

The linear high frequency emphasis filters and the LAGBC schemes were found to be very effective for contrast enhancement. Moreover, the simple recursive filter formulations developed are not only effective but result in CCD analog implementations which are much simpler than equivalent nonrecursive approaches. On the other hand, histogram modification techniques (both full frame and local area) were found to be not as consistently useful as the spatial filters for enhancing FLIR imagery.

Detailed analyses of the above contrast enhancement algorithms were given in the interim reports [1, 2]. Here we summarize the discussion of the algorithms and follow with a discussion of their effectiveness, based on their performance with imagery statistics.

High Frequency Emphasis Spatial Filters

Several forms of high-spatial frequency emphasis filters were investigated for effectiveness in crispening thermal imagery and potential for real time implementation with CCDs. High frequency emphasis filtering of thermal imagery has a twofold purpose described below.

- Attenuating the low spatial frequencies (slowly varying components) reduces the global dynamic range.
- Emphasizing the high frequencies increases local contrast, crispens edges and enhances other detail above the contrast sensitivity threshold.

Figures 23(a) and (b) show the basic high frequency emphasis filter structure employed and the corresponding frequency response shape. A low pass filter is used to derive the high emphasis filter. High frequency emphasis is obtained by a linear combination of the input and the low passed signal. This, incidentally, is similar in principle to unsharp masking often used in photography. It is possible to perform this filtering in the frequency domain taking the two-dimensional FFT and modifying the frequency coefficients and inverse transforming. We will defer comparison of FFT convolution and direct spatial convolution from an implementation view point until later, where we will see that the spatial convolution approach shows much greater promise of implementation on real time FLIR data than the FFT approach. Accordingly, two approaches to spatial domain filtering were investigated. They were two-dimensional nonrecursive and recursive filters.

- A nonrecursive two-dimensional Gaussian filter with frequency response

$$H(f_x, f_y) = \exp - \left[\frac{f_x^2 + f_y^2}{2\sigma^2} \right]$$

and impulse response

$$h(x, y) = \frac{1}{2\pi\sigma'^2} \exp - \left[\frac{x^2 + y^2}{2\sigma'^2} \right]$$

where

$$\sigma' = \frac{1}{2\pi\sigma}$$

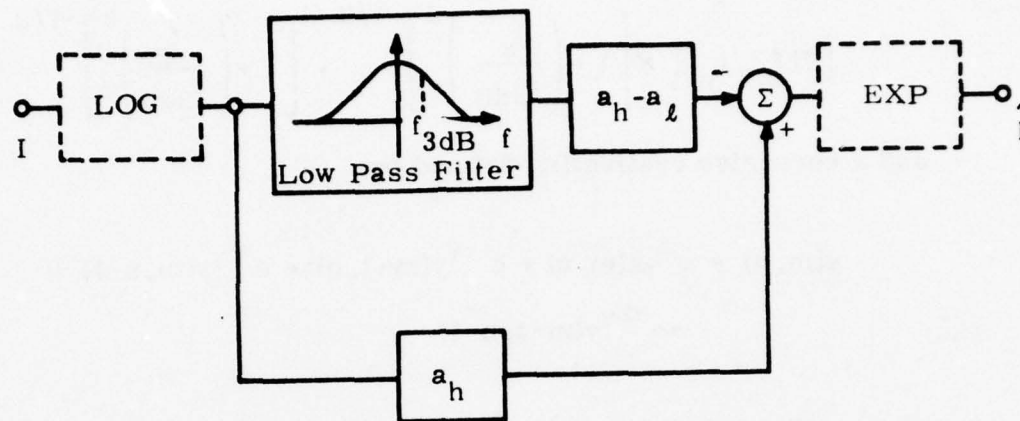


Figure 23(a). High Emphasis Filter using a Basic Low Pass Filter

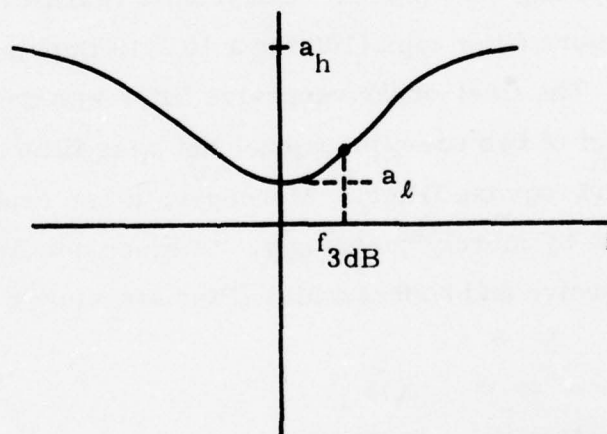


Figure 23(b). Desired High Frequency Response

- A two dimensional recursive first order filter with the frequency response

$$|H(f_x, f_y)| = \left[1 + \left(\frac{f_x}{f_{3dB}} \right)^2 \right]^{-1/2} \cdot \left[1 + \left(\frac{f_y}{f_{3dB}} \right)^2 \right]^{-1/2}$$

and a recursive realization defined by:

$$y(m, n) = \gamma^2 x(m, n) + e^{-\gamma} y(m-1, n) + e^{-\gamma} y(m, n-1) - e^{-2\gamma} y(m-1, n-1)$$

where γ is a constant, a function of the 3dB cutoff frequency of the filter.

Both these filters are separable and thus ease design and implementation. The nonrecursive filter has the advantage of being isotropic (circular symmetry) and having zero phase. But from a realization standpoint, it requires many more filter taps (100 for a 10 x 10 impulse response) and 10 line delays. The first-order recursive filter was designed as a separable product of two one-dimensional low pass filters and requires only one line delay and four tap filter. Moreover, it can realize any size impulse response by merely changing γ . A more detailed tradeoff analyses of recursive and nonrecursive filter structures can be found in Section X.

Several NVL supplied thermal (FLIR and thermoscope) images were enhanced by the two high emphasis filters. Figure 24(a) is the original thermal image and Figure 24(b) and (c) are the high emphasis filtered versions using the nonrecursive and recursive filters, respectively. The



Figure 24(a). Original Thermal Image



Figure 24(b). High Frequency Emphasis Non-recursive Filter



Figure 24(c). High Frequency Emphasis Recursive Filter

filter parameters were chosen to be given the best image visually and to optimize the contrast statistics measured on the targets. For the non-recursive filter, the size was $M = 11$; for the recursive filter the cutoff frequency was $f_c = 1/20 f_s$. For both filters, the low frequency gain was $a_l = 0.5$ and the high frequency gain was $a_h = 2.5$. Another set of images, an original that is similarly enhanced by high emphasis filtering, appears in Figures 25(a), (b) and (c). Note that the overall dynamic range has been reduced in both high frequency emphasized versions while the local contrast (especially the target detail) has been enhanced. The recursive filter did as well as the nonrecursive filter.

In summary,

- High frequency emphasis significantly improves the local contrast and compresses the scene dynamic range.
- The simple recursive filter does as well as the nonrecursive filter, which is more computationally expensive and difficult to implement, and is thus a prime candidate for implementation in future imagers.
- Although these high emphasis filters were not specifically designed to remove the MTF-induced blur in these images, they appear to improve the resolution of the targets. As such, in a limited sense, high emphasis filtering may serve to resolve the targets in the absence of a more sophisticated resolution restoration module in the imager.

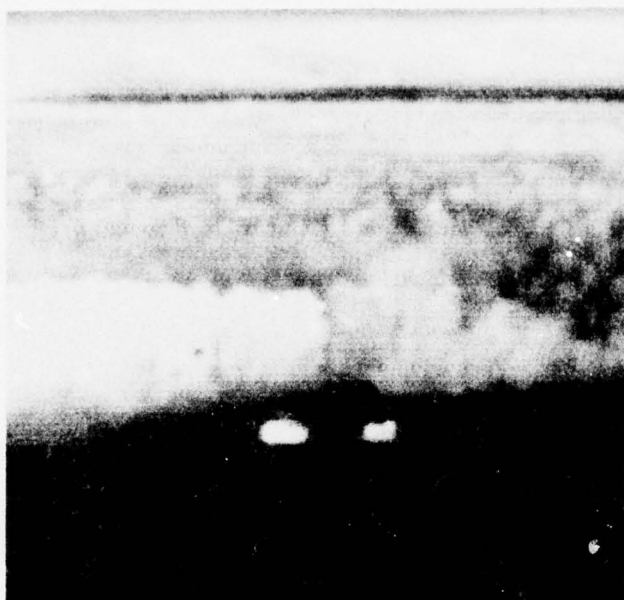


Figure 25(a). Original Thermal Image

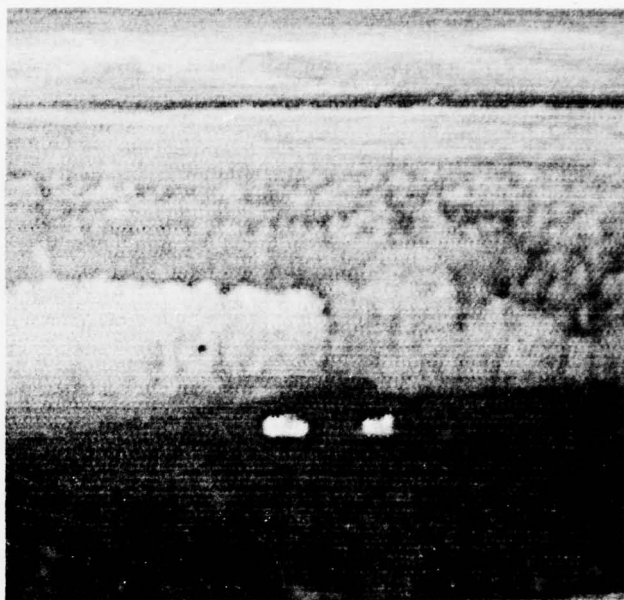


Figure 25(b). High Frequency Emphasis Non-recursive Filter

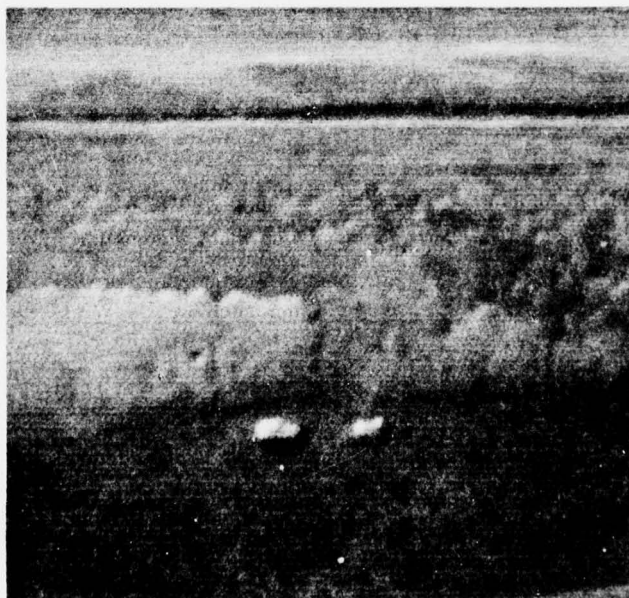


Figure 25(c). High Frequency Emphasis Recursive Filter

Local Area Gain/Brightness Control

Local area gain/brightness control (LAGBC) schemes have been proposed to locally control the gain and bias of the video so that the full luminance range of the display is used to display local information. Ideally such a scheme should:

- Vary the local average brightness (bias) so that overall dynamic range of scene is compressed;
- Enhance local variations above the contrast sensitivity threshold of the human eye; and
- Automatically fit the intensity extremes in the enhanced scene to the display limits.

The image intensity at each point is transformed based on local area statistics (the local mean M_{ij} and the local standard deviation σ_{ij} computed on a local area surrounding the point). The transformed intensity is then

$$\hat{I}_{ij} = G_{ij}[I_{ij} - M_{ij}] + M_{ij}$$

where, the local gain

$$G_{ij} = \alpha \frac{M_{ij}}{\sigma_{ij}}, \quad 0 < \alpha < 1.$$

The locally transformed intensities \hat{I}_{ij} are further scaled to the display by

$$\hat{\hat{I}}_{ij} = A_G \cdot \hat{I}_{ij} + B_G$$

where A_G and B_G are global gain and bias computed from the minimum and maximum values of \hat{I}_{ij} from the previous frame.

The LAGBC equation merely amplifies the local intensity variation around the local mean M_{ij} . The local gain G_{ij} is itself locally adaptive, being proportional to M_{ij} , to satisfy the psychovisual considerations. It is also inversely proportional to σ_{ij} so that areas with small local variance receive larger gain.

From an implementation point of view, the scheme as it stands is complex because of the need to compute the local area means and standard deviations over a sliding window centered at each point, (for a 10 x 10 window, this implies a 100-tap CCD filter). This was solved by two novel approaches to computing the local area mean and standard derivation. They were: 1) A nonrecursive approach that computed the local area statistics on nonoverlapping areas and used bilinear interpolation to obtain the statistics at all other points in the image; and 2) use of two-dimensional low pass filters to compute local means and standard derivations. The details can be found in the interim reports [1, 2]. (See also Figure 26.)

Both approaches proved to be very efficient when implemented on the computer, but the recursive filter approach is more promising from a real time implementation view point. This is because, in the interpolation approach, although the number of multiplies and adds per pixel is substantially reduced, the storing, addressing and updating of the local area statistics requires a complex digital architecture. Direct spatial filtering, on the other hand, can be performed directly with simple CCD structures as we will see in Section X.

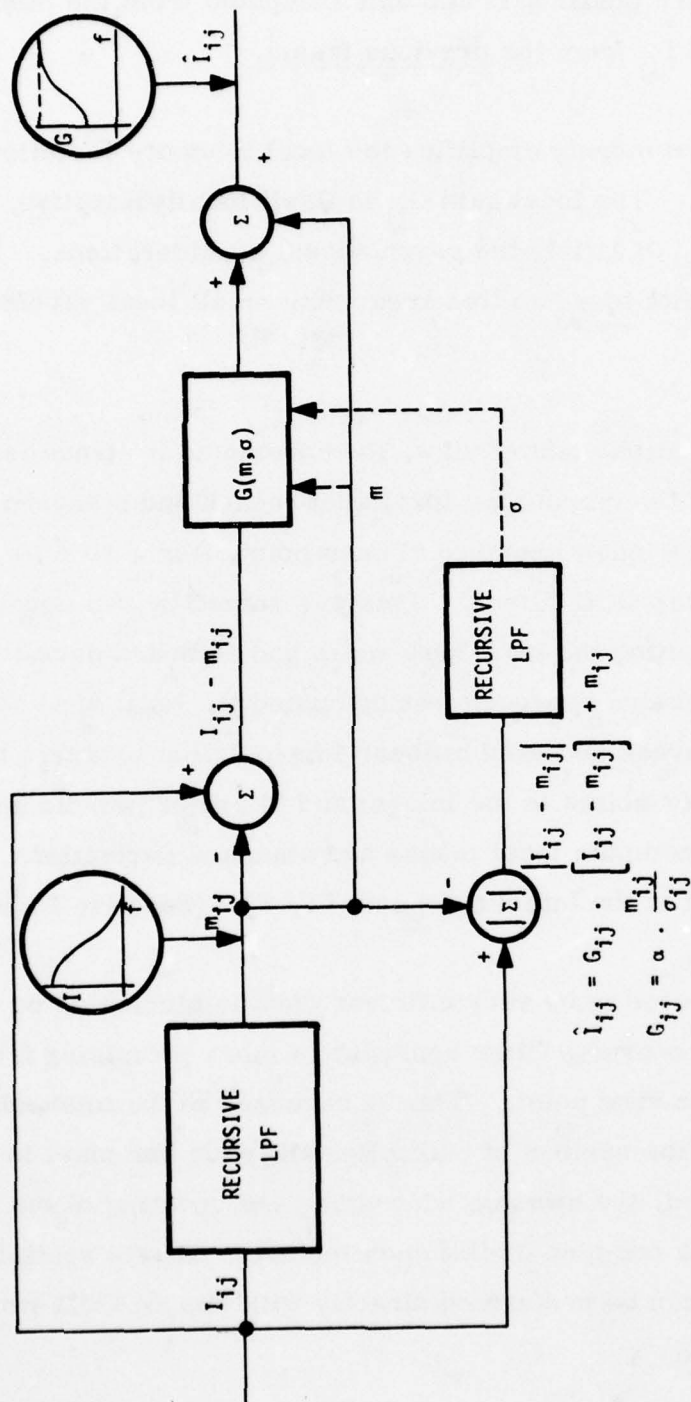


Figure 26. A Recursive Filter Approach to Local Area Gain/Brightness Control

Figures 27 through 32 show several examples of FLIR imagery that were enhanced by the recursive and nonrecursive approaches outlined above. From these, we see that each local area in the image is contrast stretched to occupy the entire display dynamic range and the recursive and nonrecursive approaches are equivalent. Not only are the targets more visible, but the details in the background have been enhanced in all of the examples shown. This, coupled with the processes being entirely automatic ("hands-off") renders the LAGBC algorithms extremely attractive for contrast enhancement. With the local area control, the FLIR operator does not have to interactively manipulate the gain and bias controls in order to selectively expand the grey-levels in some portion of the scene to search it. The following parameters were used in these examples: Minimum local gain, $G_{\min} = 2$; maximum local gain, $G_{\max} = 10$; and $\alpha = 0.3$. For the nonrecursive example, the size of the local area $M = 10$, and the corresponding cutoff frequency of the low pass filters used in the recursive filter implementation is $f_c/f_s = 1/15$.

We showed in the interim reports [1, 2] that the LAGBC formulation above, although heuristic in nature, is really adaptive high frequency emphasis filtering. The only difference between LAGBC and high frequency emphasis filtering is that the high frequency gain is not constant for all points in the image in LAGBC, but varies according to the local imagery statistics (mean and variance). Additionally, the first order low pass filters used in the recursive LAGBC formulation make real time implementation very simple because all the advantages (one line delay and the simple structure) of the recursive filters carry over. Therefore, it is clear that the above LAGBC schemes are supersets of the high frequency emphasis algorithms. As a result, we concentrated mainly on the LAGBC schemes in the implementation studies.

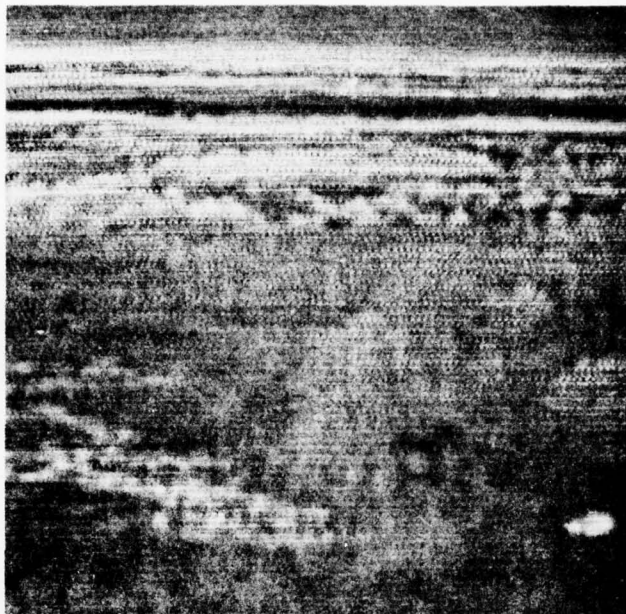


Figure 27(a). Nonrecursive Local Area Gain/Brightness Control
Transformed Version of Figure 24(a)

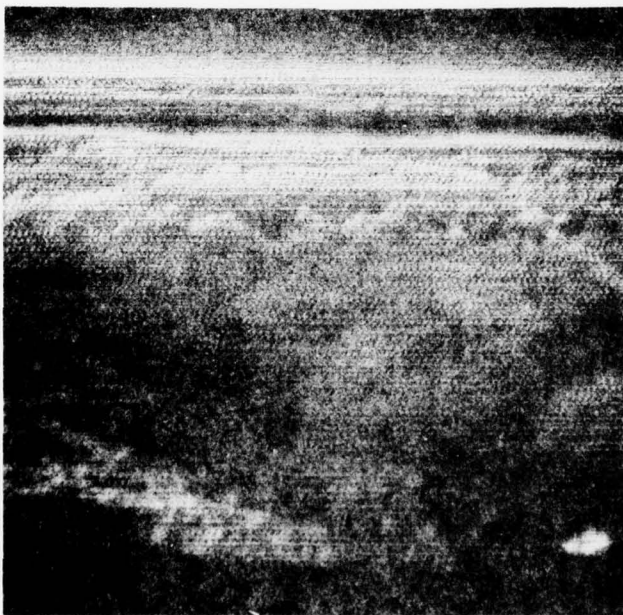


Figure 27(b). Recursive Local Area Gain/Brightness Control
Transformed Version of Figure 24(a)

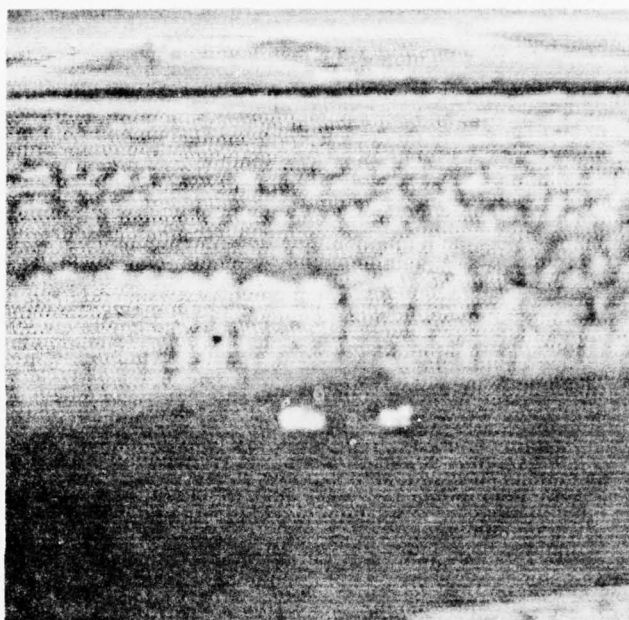


Figure 28(a). Nonrecursive Local Area Gain/Brightness Control
Transformed Version of Figure 25(a)

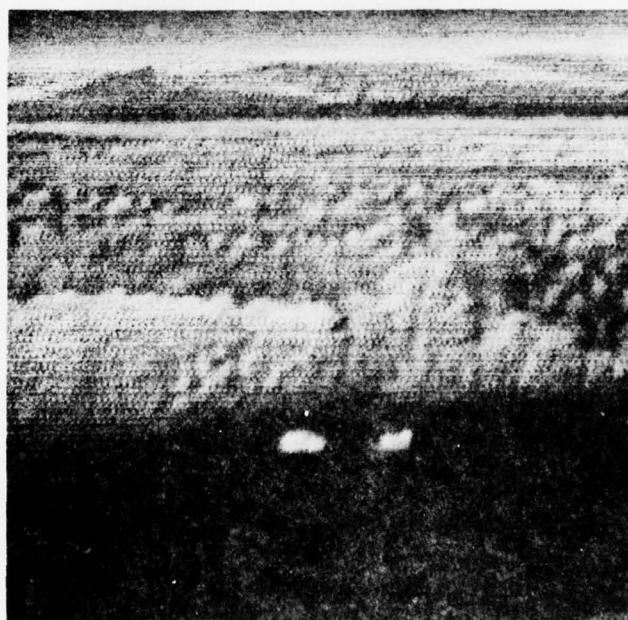


Figure 28(b). Recursive Local Area Gain/Brightness Control
Transformed Version of Figure 25(a)

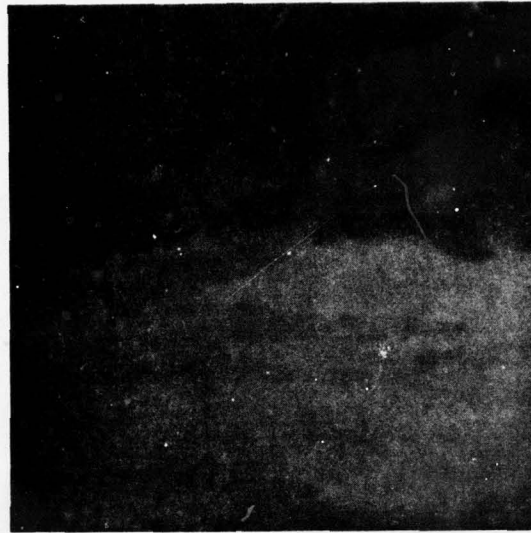


Figure 29(a). Original Thermal Image



Figure 29(b). Recursive LAGBC Enhanced

AD-A049 564

HONEYWELL INC MINNEAPOLIS MINN SYSTEMS AND RESEARCH --ETC F/G 17/5
AUTOMATED IMAGE ENHANCEMENT TECHNIQUES FOR SECOND GENERATION FL--ETC(U)
DEC 77

DAA653-76-C-0195

UNCLASSIFIED

77SRC93

NL

2 OF 4
AD
A049564



2

OF

4

AD

A049564



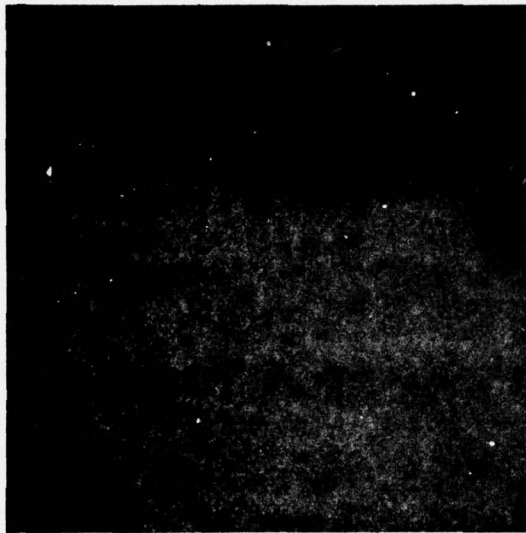


Figure 30(a). Original Thermal Image



Figure 30(b). Recursive LAGBC Enhanced

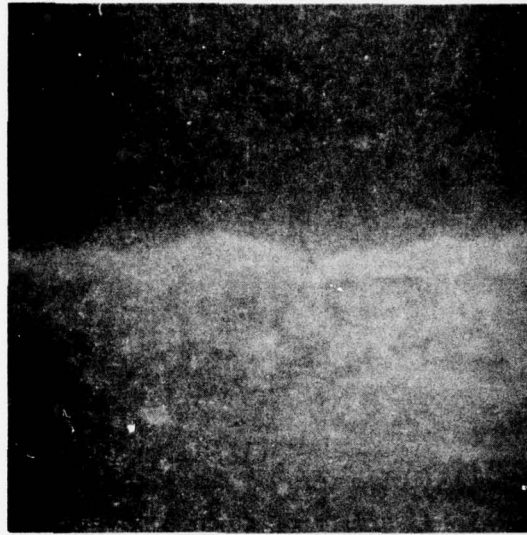


Figure 31(a). Original Thermal Image

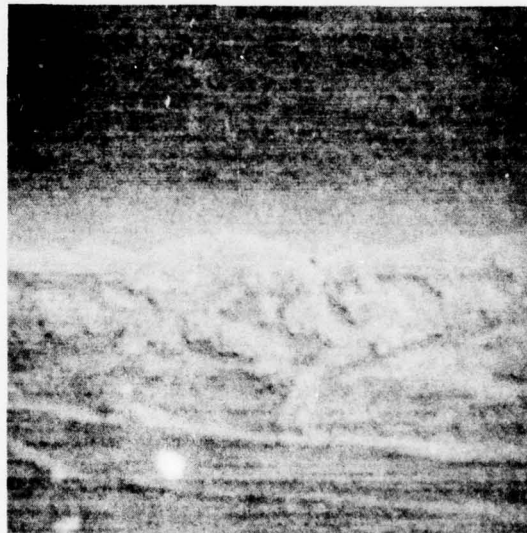


Figure 31(b). Recursive LAGBC Enhanced

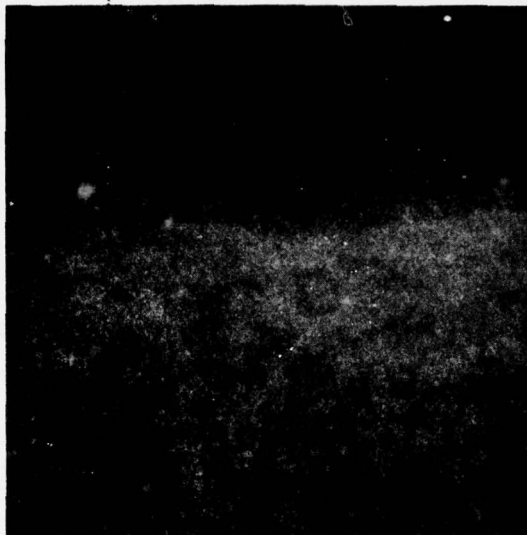


Figure 32(a). Original Thermal Image



Figure 32(b). Recursive LAGBC Enhanced

Homomorphic Filters--Homomorphic filtering was included in the high frequency filters as implied by the LOG and EXP boxes of Figure 23(a). Homomorphic filtering assumes a multiplicative model of the image, which is indeed true of the visible and near IR parts of the EM spectrum where the image can be modeled as a product of low frequency illumination and high frequency reflectivity. This model is not directly applicable to thermal IR imagery. Moreover, experiments with homomorphic filtering for thermal images showed similar results to that of linear filtering, but the results were extremely sensitive to the high and low frequency gains chosen (owing to the nonlinearity of the process). Further, since the target intensities can be in any part of the grey scale, nonlinear processing might unfavorably affect target visibility in some cases. As a result, we chose not to investigate homomorphic filtering further.

Histogram Modification Techniques--Some grey levels occur more often than others in an image. Sometimes it is desirable to make more effective use of the available grey levels. Histogram modification techniques nonlinearly map the intensities in the original image to another domain so that the new intensities have a specified distribution. When the new distribution is uniform (every grey level occurs equally often), the mapping is known as histogram equalization. The following histogram modification schemes were investigated for enhancement of thermal imagery. (See the first interim report [1].)

1. Full Frame Histogram Equalization--Figure 33 is a histogram-equalized version of Figure 27(a) (which has been subjected to LAGBC) and serves to illustrate the main drawbacks of grey-level mapping schemes in general. Histogram equalization maximizes the first

order entropy (an information measure) of the grey-level distribution of the image, but takes no spatial information into account. The target detail in Figure 27(a) was in the upper "tail" of the histogram. Equalization compresses the less frequent grey levels together--in this case, those representing the target details--and the corresponding image loses all the target detail. Much undesirable texture (such as banding) has been amplified, however.

2. Histogram Hyperbolization--This is based on the premise that the perceived brightness is roughly proportional to the logarithm of the display intensity I . The observed brightness levels ($\log I$) should therefore be uniformly distributed. To achieve this, the distribution of the displayed intensities should be hyperbolic, instead of uniform. This was found to be somewhat more pleasing than

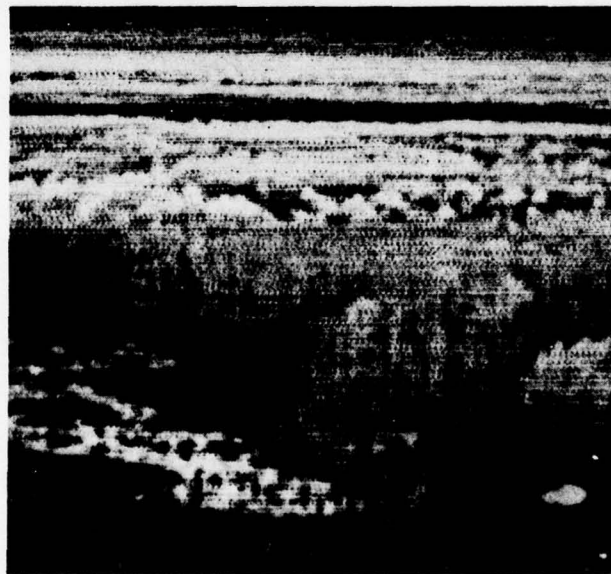


Figure 33. Histogram Equalized Version of the LAGBC Image in Figure 27(a).

equalization, but the target details were lost here also, and the same objections raised to equalization still apply.

3. Local Area Histogram Equalization (LAHE)--Several local area histogram equalization schemes were tested with FLIR imagery [1]. Here, grey levels in a small area (window) of the image were modified using a transformation computed from the grey-level histogram of a larger subregion surrounding the local area. It was found, however, that LAHE did not significantly increase target resolution or details, but accentuated severe banding and clutter in the images. Besides being ineffectual for tactical FLIR image enhancement, the LAHE techniques are also very complex to implement, especially with overlapping local areas.

STATISTICAL ANALYSIS OF CONTRAST ENHANCED IMAGERY

As discussed above, the following contrast enhancement (CE) schemes were computer simulated and used to enhance thermal imagery in order to evaluate their effectiveness.

- High frequency emphasis recursive filter
- High frequency emphasis nonrecursive filter
- LAGBC nonrecursive realization
- LAGBC recursive realization
- Local area histogram modification scheme
- Global histogram modification

From the results of simulation on a small set of images, the best set of parameters were selected for each algorithm [1]. Also based upon these simulations, we selected a subset of three algorithms--the recursive high frequency emphasis filter, and the recursive and nonrecursive local area gain brightness algorithms--to enhance a much larger set (40) of FLIR images. The resultant (120) enhanced images served a twofold purpose: 1) They were input to the human factors evaluation process; and 2) they served a statistical base of imagery statistics which were used to quantify the performance of the three contrast enhancement algorithms. In this subsection we will summarize the results of this statistical analysis.

Contrast and Intensity Statistics

Figure 34(a), (b) and (c) are plots of the mean intensity (target and backgrounds) before (X axis) and after (Y axis) the three respective contrast enhancements. We found this form of display useful to represent changes in an image feature before and after each enhancement. Note that if there were no change, all the points would be on the 45° (unity slope) line on these plots (solid line). The scatter around this line indicates how the enhanced statistics have departed from the original. With reference to Figure 34(a), (b) and (c), we note that:

1. All three CE schemes change the mean levels of backgrounds and targets substantially.
2. The two LAGBC schemes (Figure 34(b) and (c)) show greater separation of target and background intensity means after (Y axis) than before (X axis) enhancement.

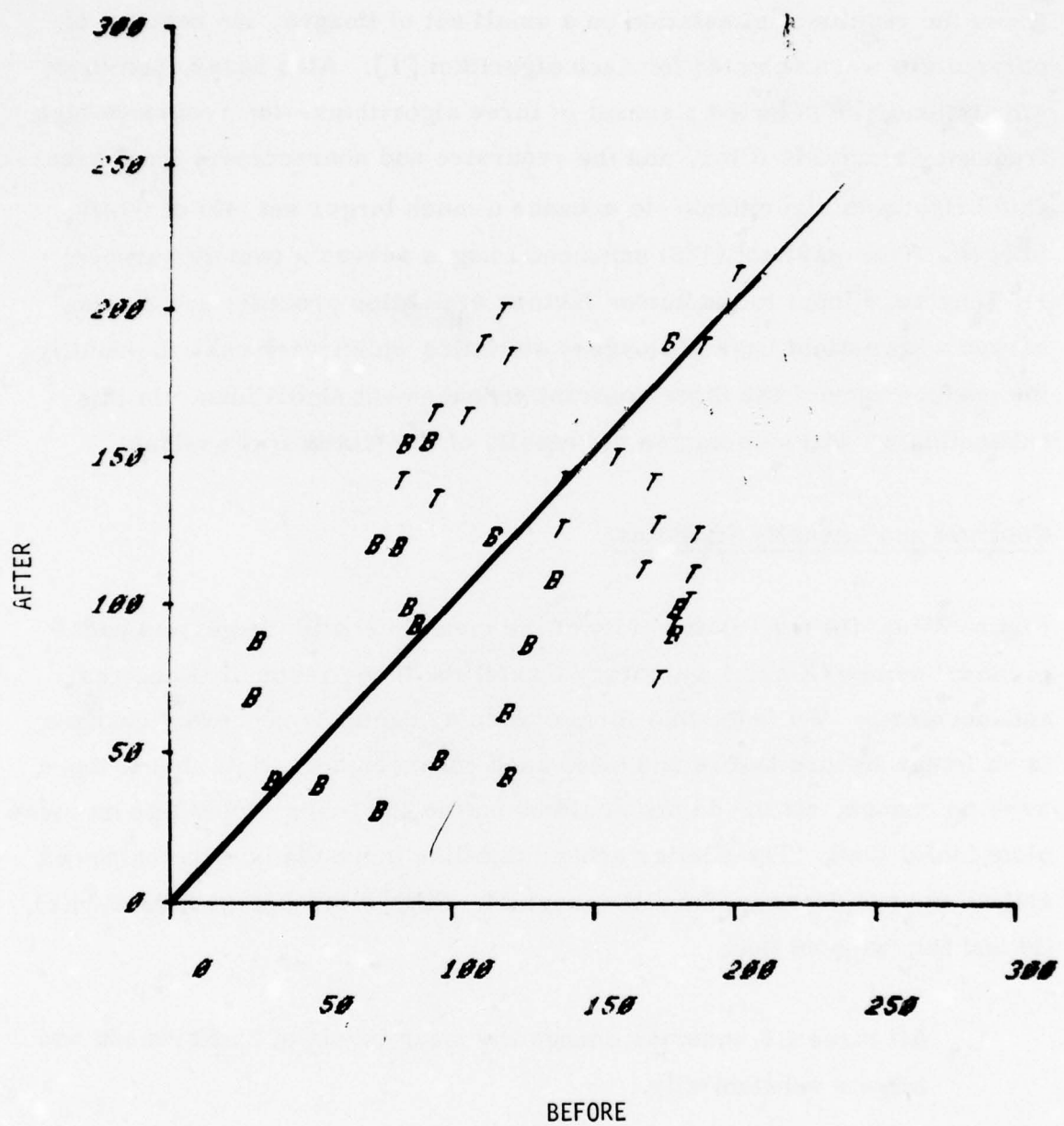


Figure 34(a). Mean Intensity of Target (T) and Background Before vs. After High Frequency Emphasis (Recursive) Filtering

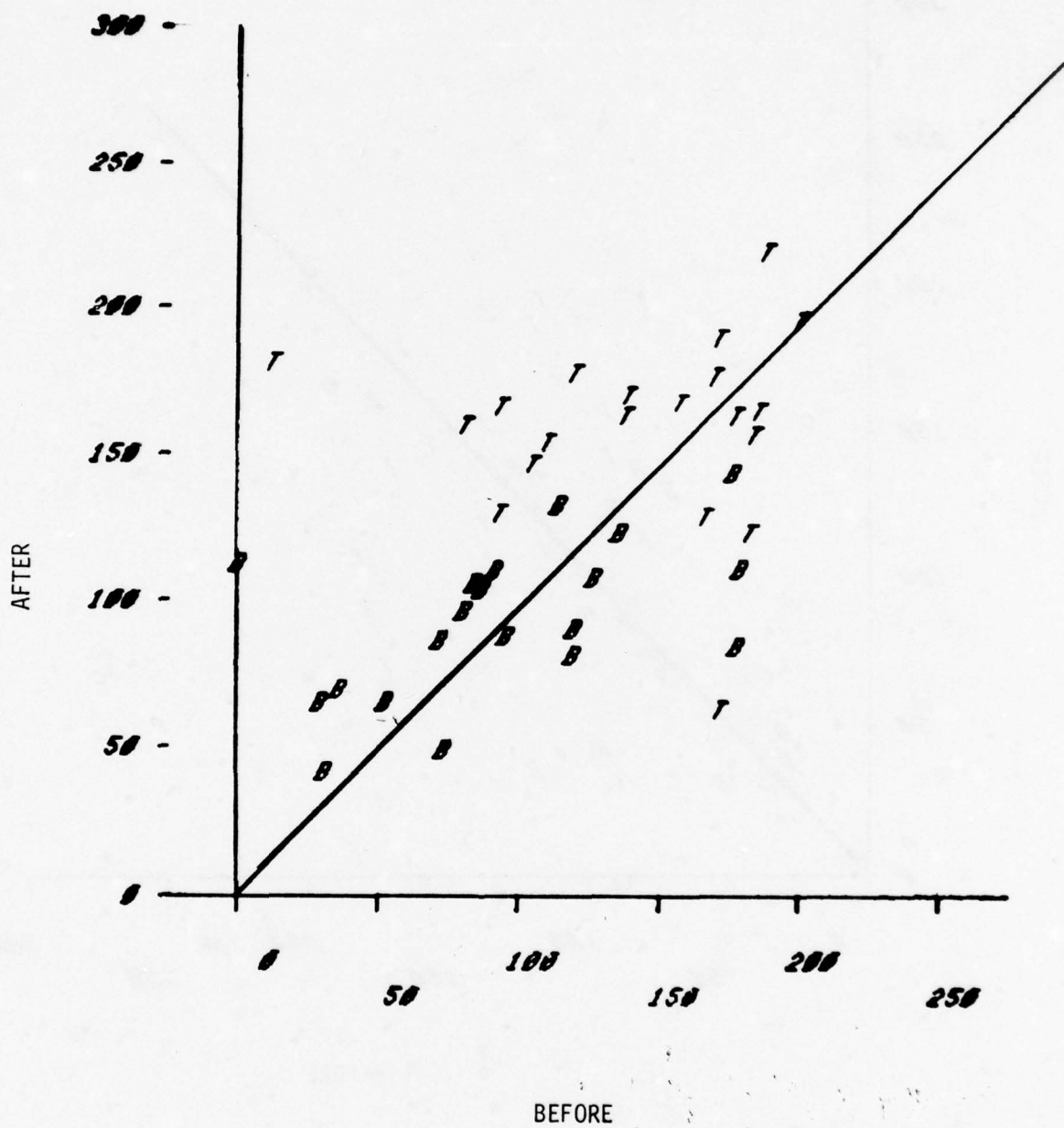


Figure 34(b). Mean Intensities of Target (T) and Background Before and After

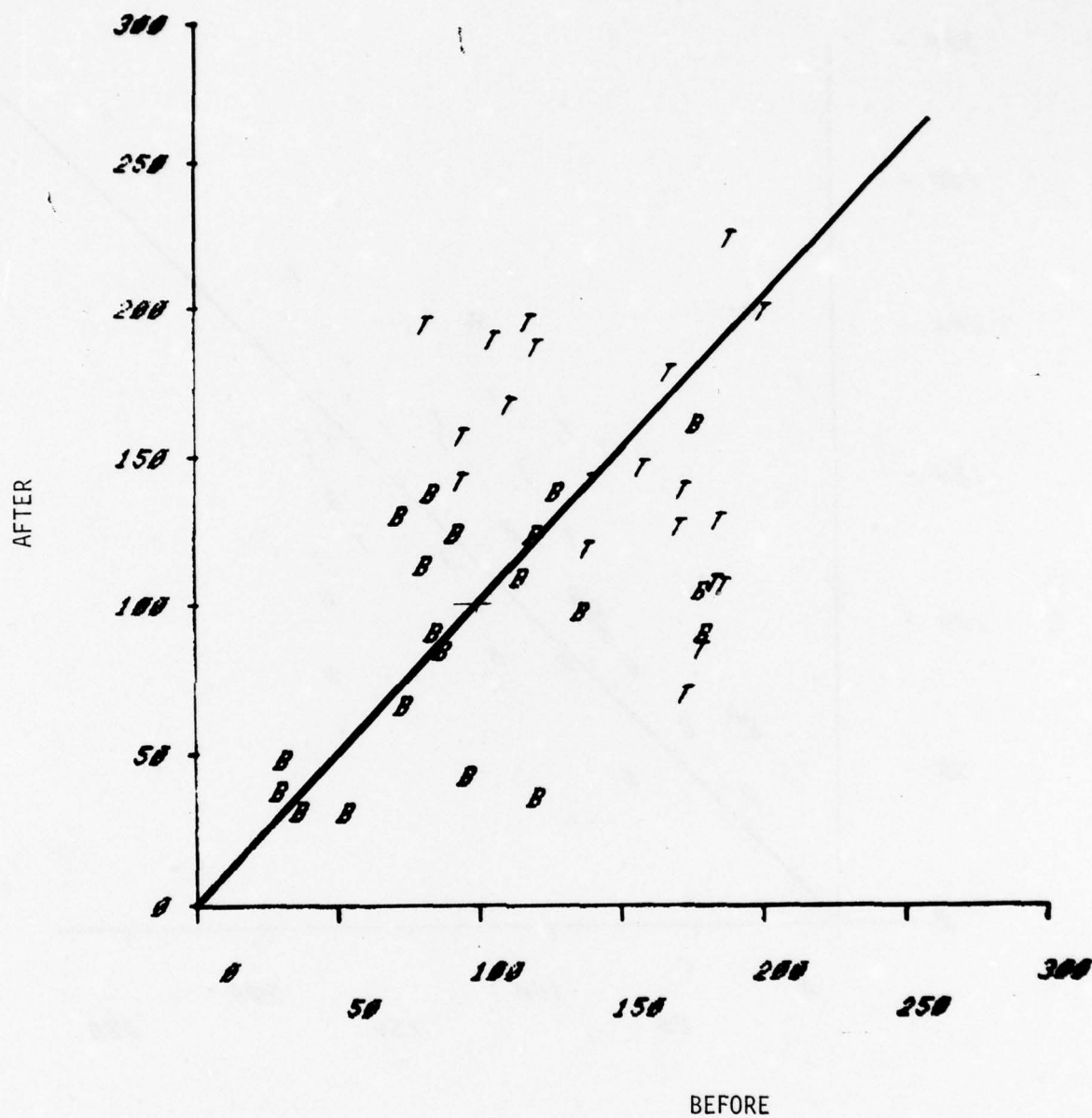


Figure 34(c). Mean Intensity of Target (T) and Background Before vs. After Local Area Gain/Brightness Control (Recursive)

Figure 35(a), (b) and (c) shows the standard deviation of background and target intensities for the three CE schemes in a format similar to the previous plots. Note that the high frequency emphasis filter and the LAGBC schemes tend to make the target and background standard deviations equal in all images, i.e., between 20 to 30, regardless of the initial standard deviations. This bears out the corresponding property of the algorithm proved in [1].

More revealing than the above are the plots of the contrast statistics (average, peak and edge) in Figures 36(a), (b) and (c). The average and peak contrasts are normalized with respect to the background. These plots depict the percentage change in the contrast (normalized with respect to the original contrast) versus original contrast. For clarity, the X (original contrast) axis is divided into a small number of discrete regions, and the percentage change in contrast is averaged in each of these regions and shown in the form of bar charts. Referring to the three curves in the Figures 36(a), (b) and (c) which correspond to the LAGBC nonrecursive, LAGBC recursive and the recursive high frequency emphasis filter, respectively, we note that in the low contrast images the CE algorithms improve the contrast 300 to 400 percent while they actually decrease the target contrast in the very high contrast images. This is exactly what they were designed to do. Moreover, the adaptive LAGBC schemes improve the contrast in low contrast images more than the nonadaptive high frequency emphasis algorithm. This is again to be expected, because the LAGBC accentuates areas of low contrast in an adaptive manner. The recursive and nonrecursive alternatives to LAGBC have performed equally well in enhancing all three contrast measures, and are superior to the high frequency emphasis filter.

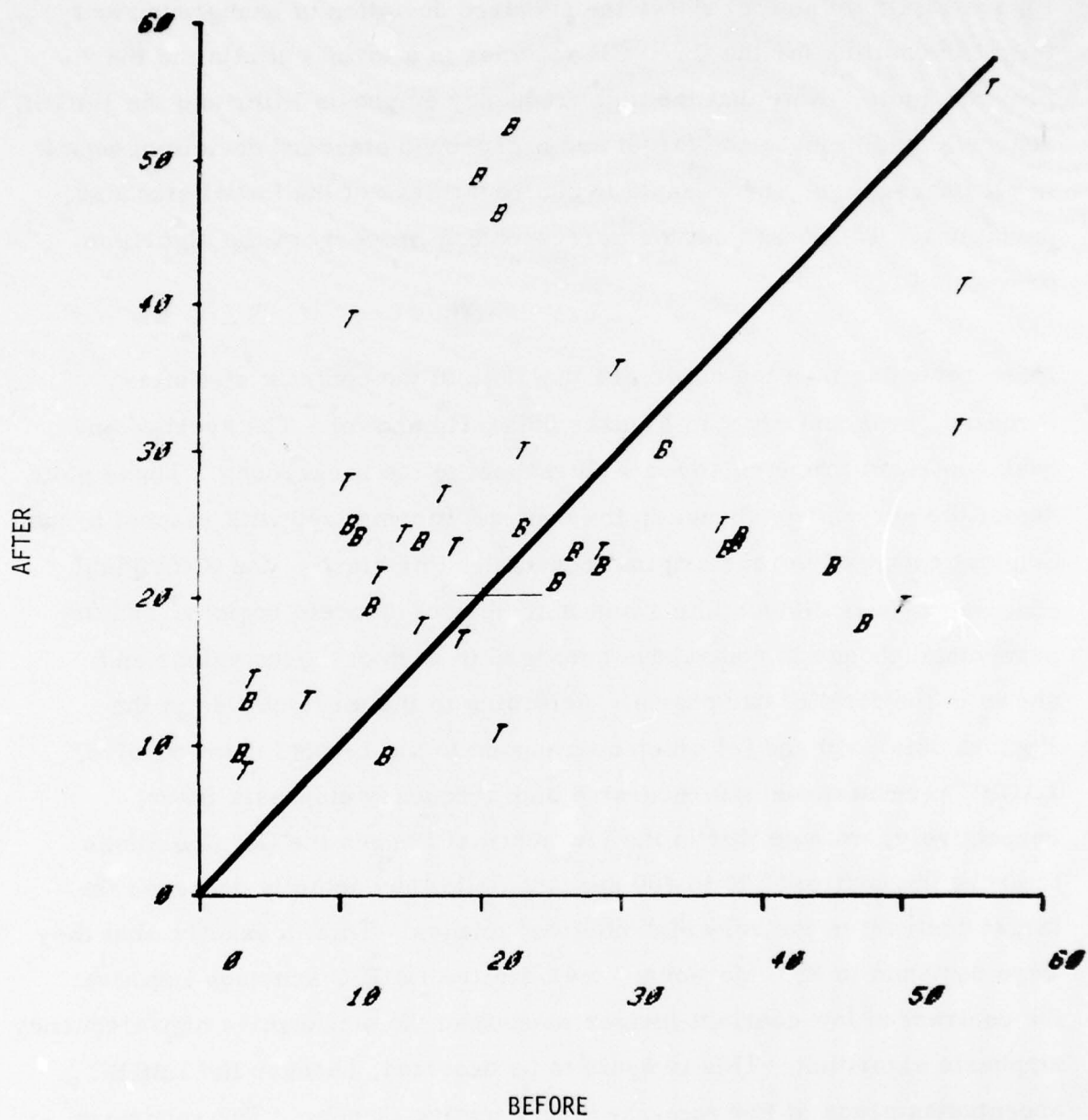


Figure 35(a). Standard Deviation of Target (T) and Background (B) Intensities Before vs. After High Frequency Emphasis Filtering (Recursive)

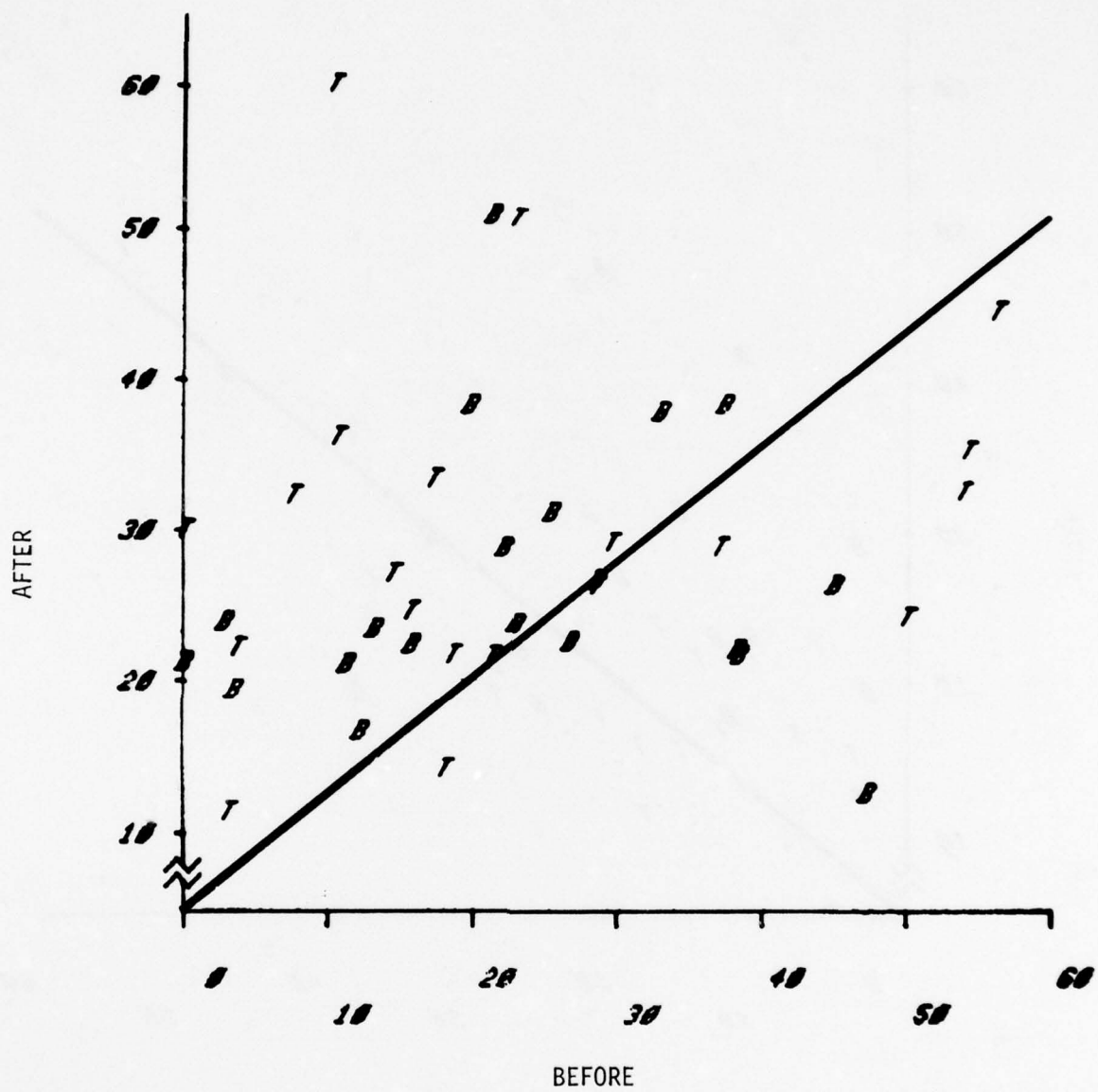


Figure 35(b). Local Area Gain/Brightness Control (Nonrecursive)

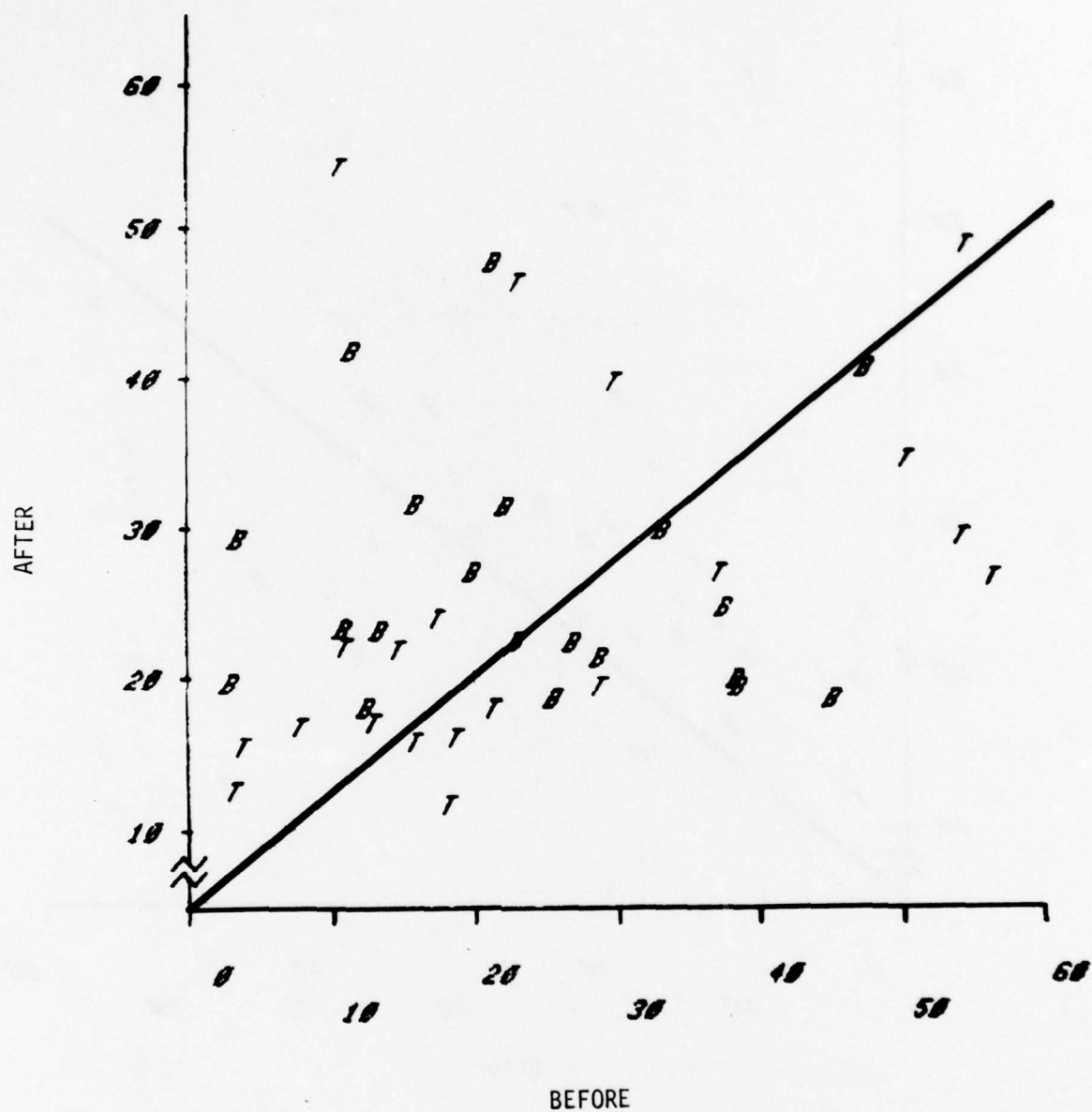


Figure 35(c). Local Area Gain/Brightness Control (Recursive)

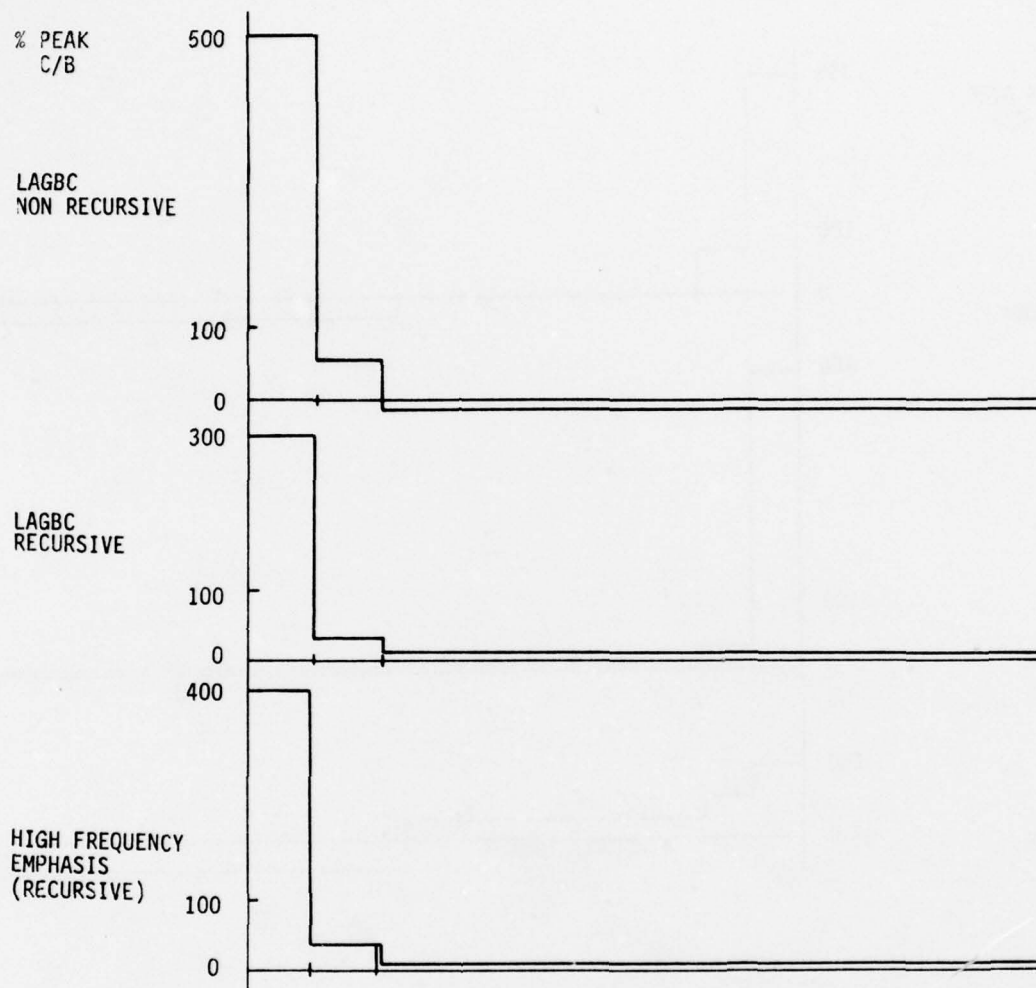


Figure 36(a). Percentage Change in Peak Contrast After Enhancement vs. Original Peak Contrast

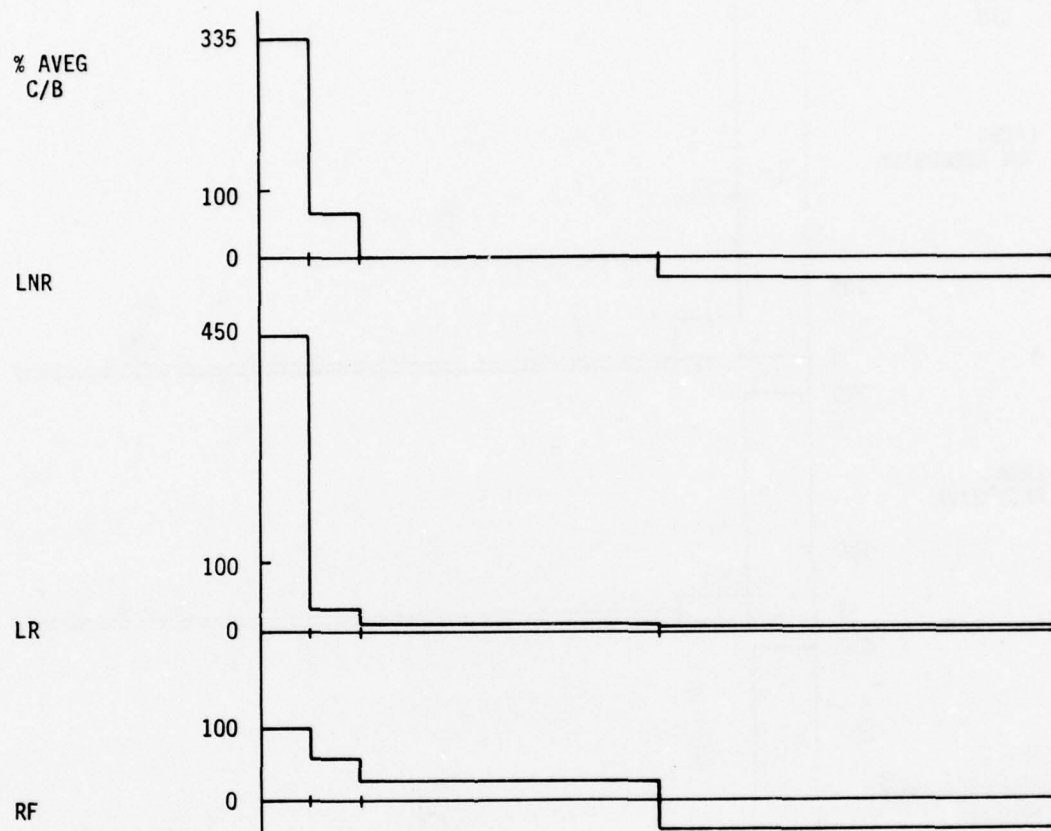


Figure 36(b). Percentage Change in Average Contrast After Enhancement vs. Original Average Contrast

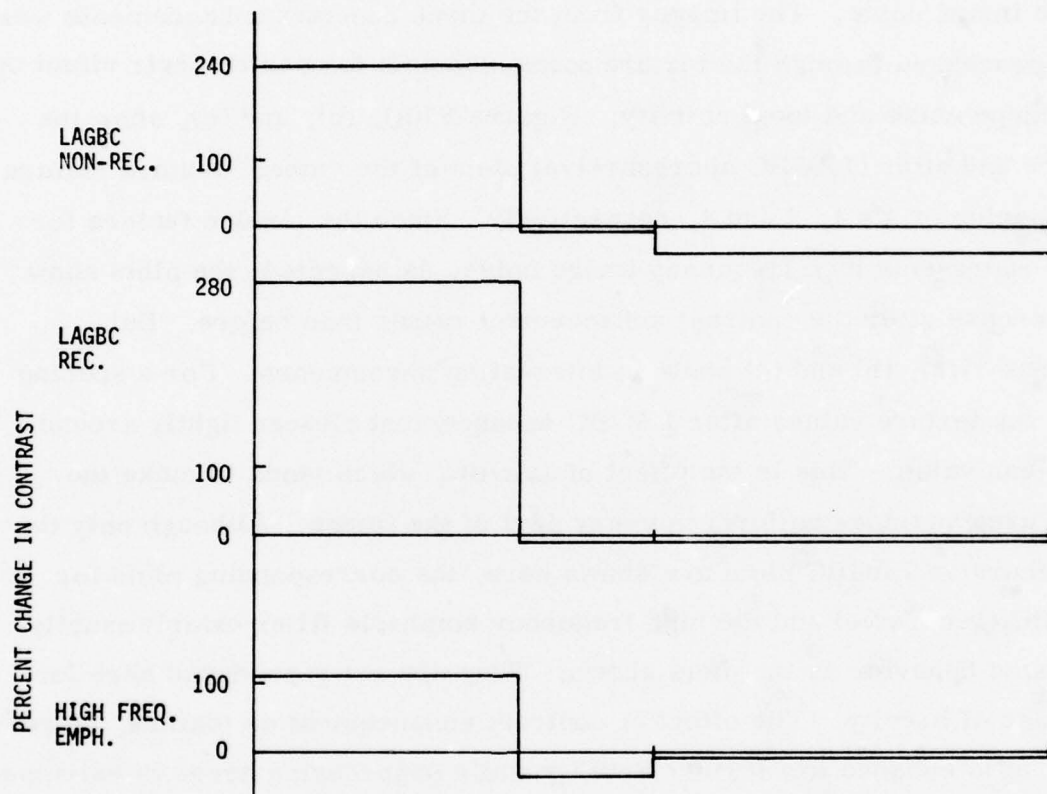


Figure 36(c). Percentage Change in Edge Contrast After Enhancement vs. Original Edge Contrast

Texture

In Section III we saw how the texture statistics were a measure of image activity (detail) and also (for a sufficiently small period Δ) a good measure of the image noise. The images from the three contrast enhancements were also processed through the texture measurements to quantify their effect on the image noise and local activity. Figures 37(a), (b), and (c), show the before and after (LAGBC nonrecursive) plots of the "mean" texture feature for spacing of $\Delta = 1, 2$ and 4 , respectively. Since the texture feature for $\Delta = 1$ represents high frequency image noise, as expected, the plots show an increase after the contrast enhancement rather than before. But Figures 37(a), (b) and (c) show an interesting phenomenon. For a spacing $\Delta = 4$ the texture values after LAGBC enhancement cluster tightly around the mean value. This is the effect of LAGBC, which tends to make the local area variance uniform in every part of the image. Although only the nonrecursive LAGBC plots are shown here, the corresponding plots for LAGBC (recursive) and the high frequency emphasis filter exhibit exactly the same behavior as the plots shown. They are not reproduced here for the sake of brevity. The effect of contrast enhancement on texture, therefore, is to enhance low texture activity while suppressing areas of extremely large texture in the image. While this is useful to compress the dynamic range in the image, the attendant distortion in the magnitude of the textural differences may result in degradation (texture dependent) of perspective perception in the image.

Contrast Enhancement Summary

A total of six algorithms were investigated for contrast enhancement, listed below.

- High frequency emphasis recursive filter
- High frequency emphasis nonrecursive filter
- LAGBC nonrecursive realization
- LAGBC recursive realization
- Full frame histogram modification
- Local area histogram modification scheme

Each of these schemes is governed by several basic parameters which were tuned for best performance. Based on an initial production run on a limited set of images, the optimum parameters were determined for each algorithm and the three most promising of these were selected for a larger calculation study. These were the recursive and nonrecursive LAGBC schemes and the recursive high frequency emphasis scheme. These three were evaluated with a larger set of FLIR images and image statistics were measured on them. Based on this analysis, the LAGBC schemes were found to be most effective in improving target contrast and edge enhancement. The recursive approach to LAGBC is the scheme more capable of implementation in real time because of the structural simplicity of the recursive filters (only one line delay, few multiplier weights). Moreover, the LAGBC schemes were shown to be really adaptive high frequency emphasis, and therefore are supersets of the high frequency emphasis schemes. Histogram equalization

and greyscale modification techniques, in general, were determined to be ineffectual for enhancing FLIR image contrast in a noninteractive environment.

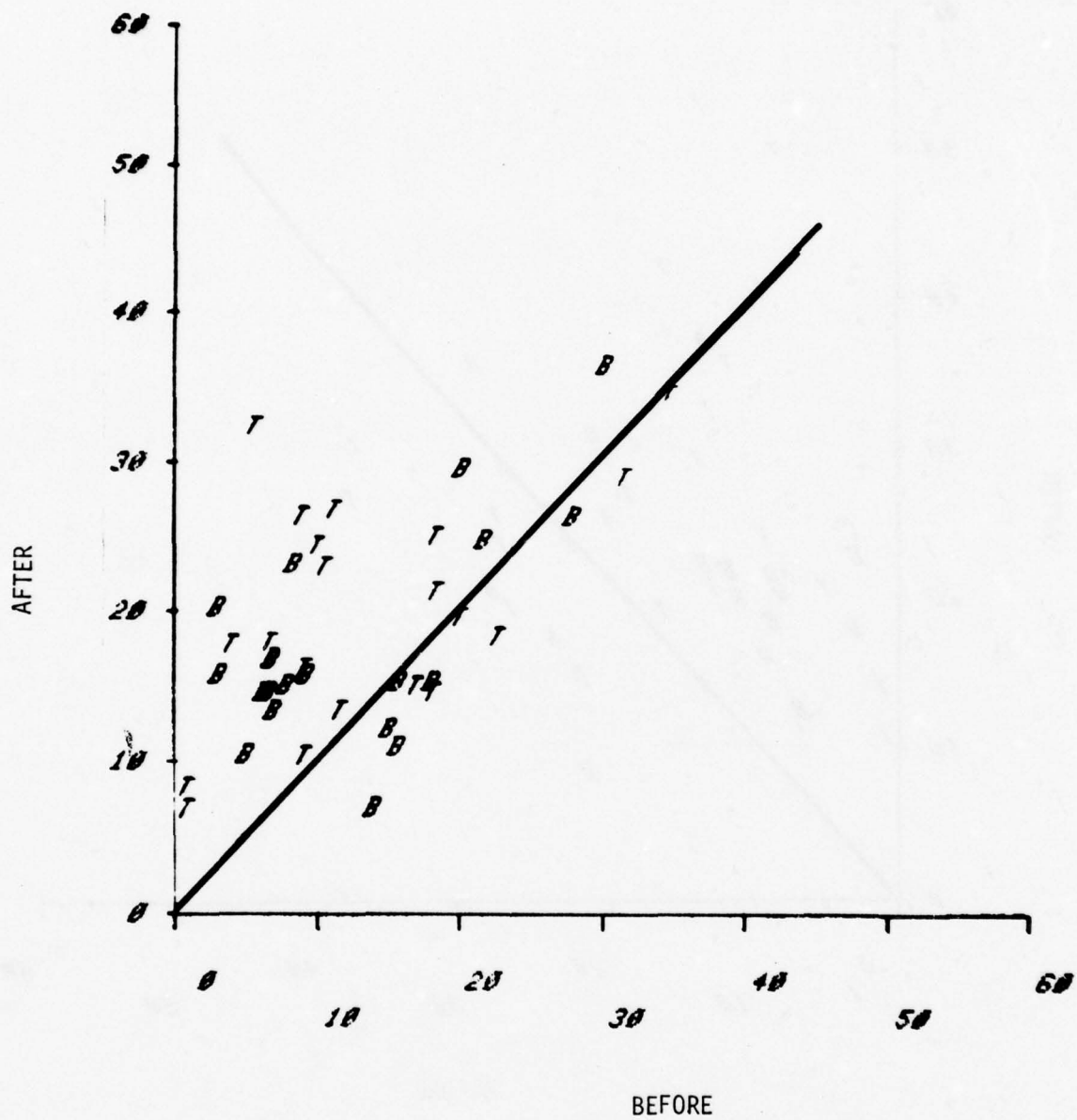


Figure 37(a). "Mean" Texture Feature ($\Delta = 1$) Before vs. After LAGBC (Nonrecursive)

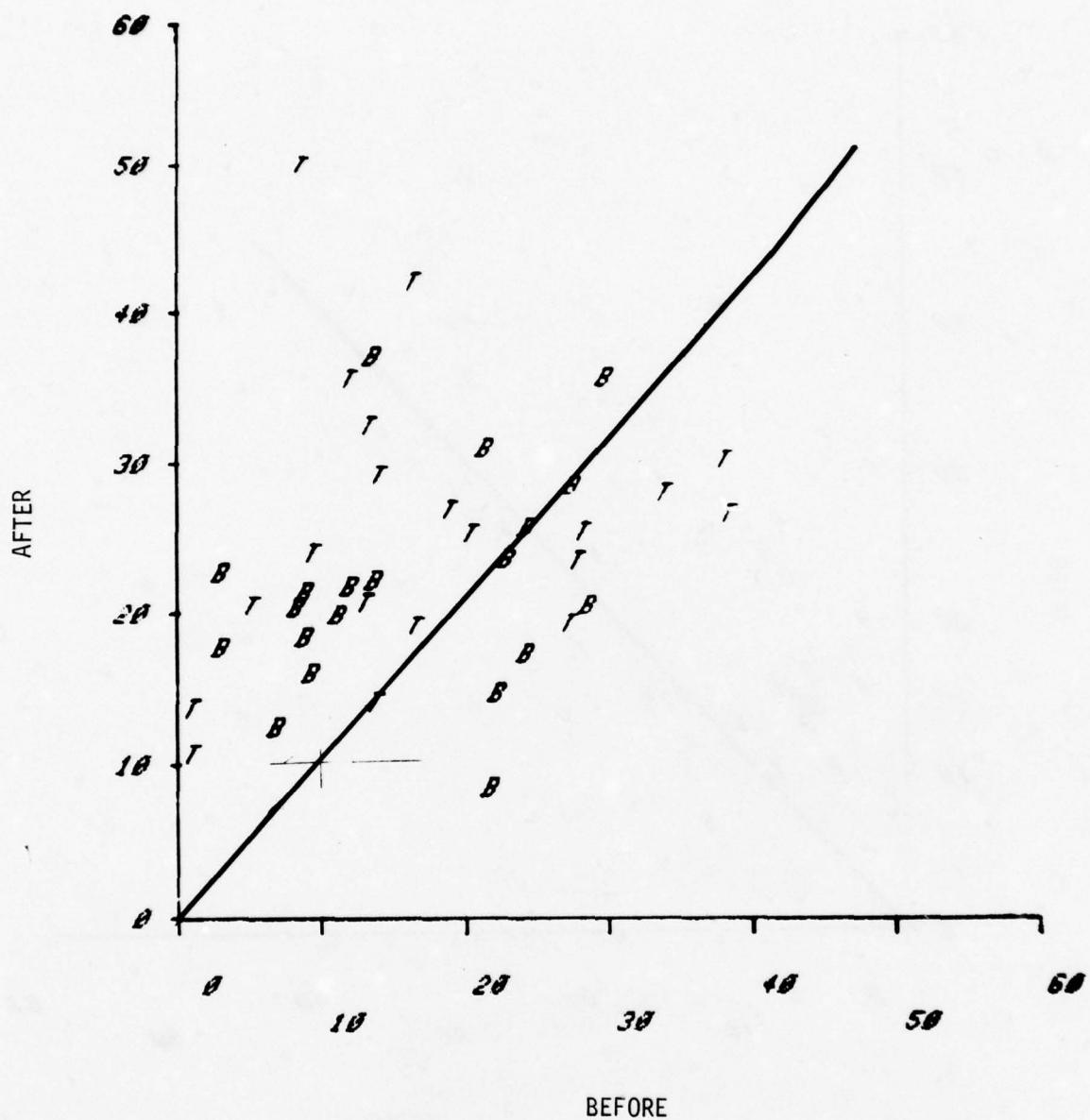


Figure 37(b). "Mean" Texture Feature ($\Delta = 2$) Before vs. After LAGBC (Nonrecursive)

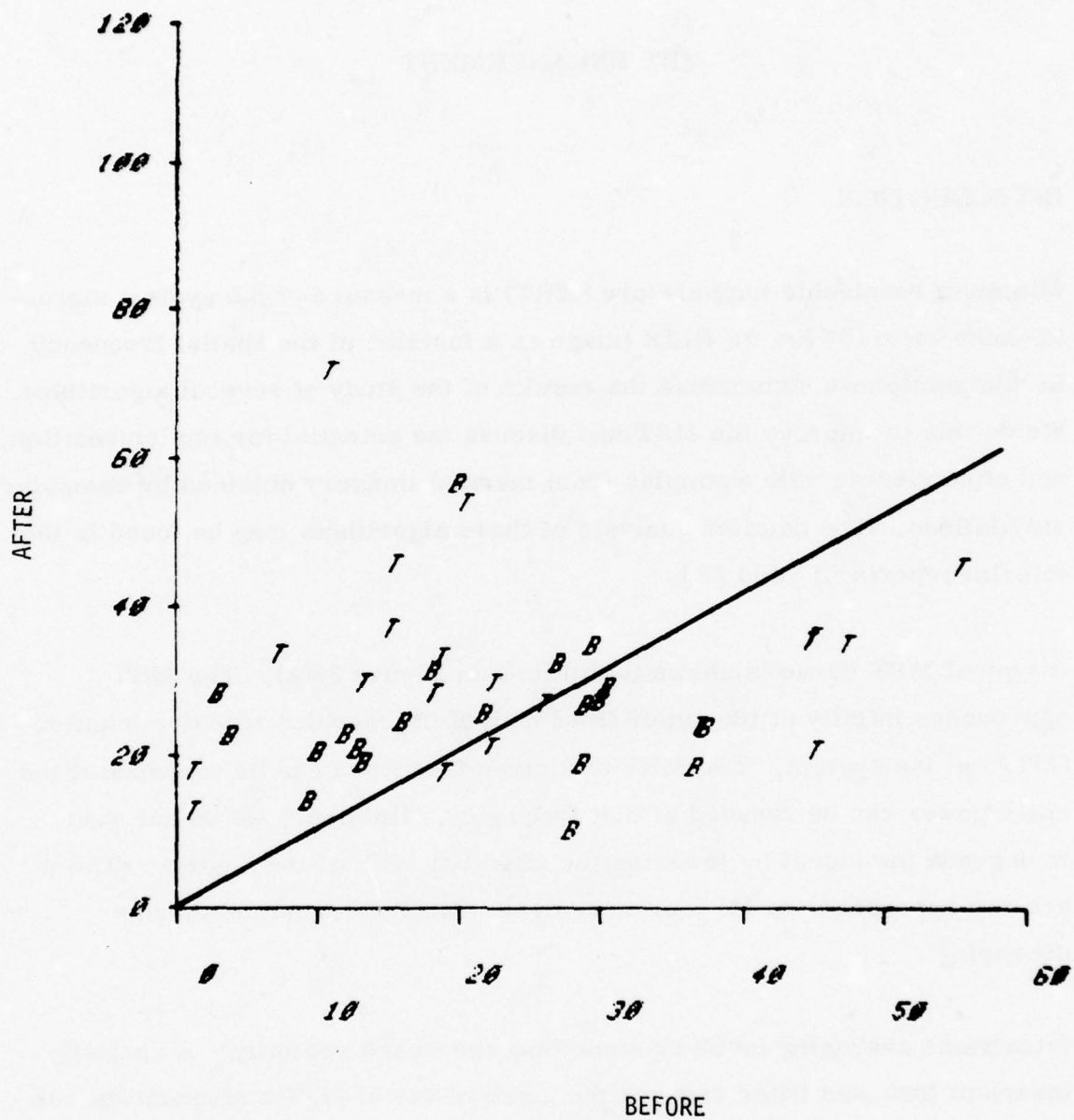


Figure 37(c). "Mean" Texture Feature ($\Delta = 4$) Before vs. After LAGBC (Nonrecursive)

SECTION V

MRT ENHANCEMENT

INTRODUCTION

Minimum resolvable temperature (MRT) is a measure of the system signal-to-noise ratio (SNR) of a FLIR image as a function of the spatial frequency. In this section we summarize the results of the study of several algorithms. We do this to improve the MRT and discuss the potential for implementation and effectiveness with examples from thermal imagery obtained by computer simulations. The detailed analysis of these algorithms may be found in the interim reports [1] and [2].

A typical MRT curve is shown (solid line) in Figure 38(a). The MRT approaches infinity at the cutoff frequency of the modulus transfer function (MTF) of the system. The MRT at a given frequency can be enhanced if the noise power can be reduced at that frequency. However, we do not want to degrade the signal by lowering the effective MTF of the system. There are two approaches to MRT enhancement: intraframe and interframe averaging.

Intraframe averaging involves smoothing the image spatially. A spatially invariant low pass filter reduces the noise power at higher frequencies but it also attenuates the higher frequency information corresponding to edges, resulting in blurred detail. This effect is shown by the dotted line in the MRT curve of Figure 38(a), where the MRT of the enhanced image approaches

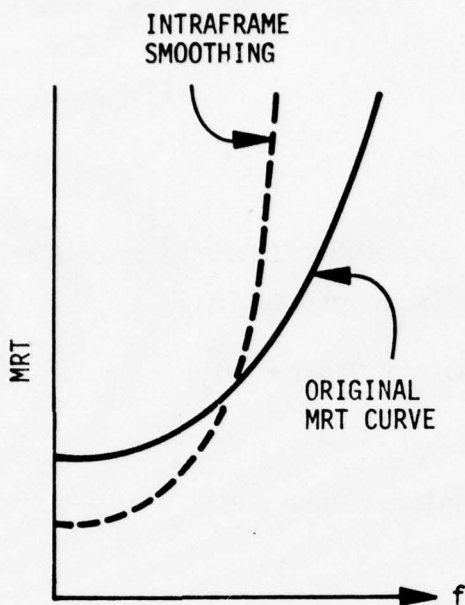


Figure 38(a). Original MRT Curve and Intra-frame Smoothed Curve

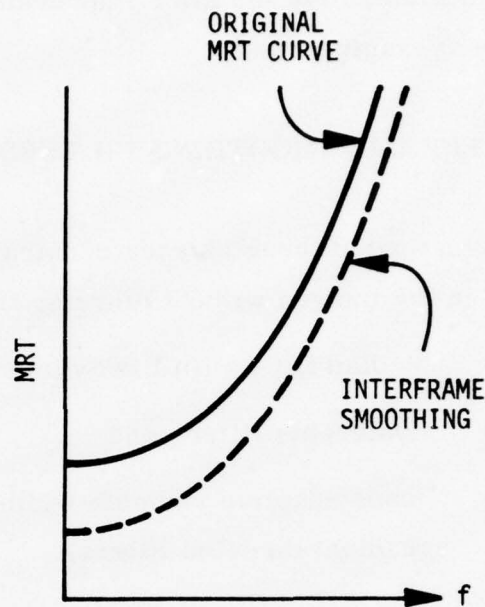


Figure 38(b). Original and Inter-frame Smoothed MRT Curves

infinity at a lower frequency due to the degradation of the system MTF by the low pass filter. We investigated algorithms for intraframe MRT enhancement that would yield higher overall SNRs while preserving edges in FLIR imagery.

Interframe MRT enhancement is temporal averaging of several registered and stacked FLIR frames. Since the noise is uncorrelated between successive frames, the noise variance is reduced by a factor n , the number of stacked frames. The limitations are that resolution is lost if registration is not maintained. The effective data rate is reduced. Figure 38(b)

(dotted line) shows the MRT improvement that can be expected from inter-frame averaging.

INTRAFRAME SMOOTHING FILTERS

We investigated three classes of intraframe smoothing filters to smooth the noise in the images without blurring the high frequency detail :

- Median filters (full two-dimensional, and separable);
- Hysteresis filter ; and
- Scene adaptive variable width filters (curvature and gradient directed filters)

Of the above, the median filters and the adaptive variable width filters proved to be the most effective for intraframe smoothing and received a great deal of our attention. Hysteresis filter, on the other hand, was not very effective and was not further investigated. (See interim reports for a discussion of this filter [1, 2].)

Median Filters

Figure 39 shows the basic structure of a 3×3 median filter; the median intensity in a small window replaces the pixel intensity at the center of the window. We have investigated 3×3 and 5×5 window median filters. The proposed advantage of the median filter is that it does not blur a step edge because the median intensity of the window is dominated by the intensities in that side of the edge the window is centered on. Corners are slightly affected, however, but this is often deemed insignificant for FLIR

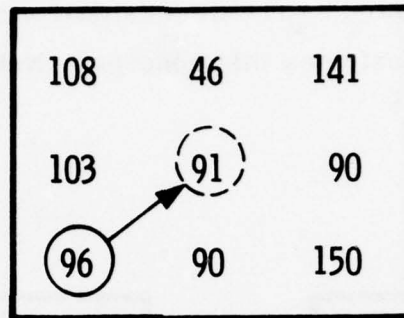


Figure 39. A 3 x 3 Median Filter

imagery. This filter is most useful with spiky noise (as a prefilter operation for edge operators, for example; and as we shall see later, as a post filter after edge emphasis filtering).

Separable Median Filter

Two-dimensional median filtering is difficult to implement in real time at videorates because it requires real time ordering (sorting of M^2 points for a $M \times M$ median filter. Analog diode networks can be used to find the median of M input voltages and are capable of operating at video rates [4]; but beyond $M = 5$ (30 diodes) they grow very complex in structure (because they require approximately $.5 M!$ diodes to implement). A nine point median network requires 630 diodes (for a 3 x 3 median filter). A 25 point

median filter requires 67×10^6 diodes! Therefore, we also investigated the separable median filter, i. e., one-dimensional M point median filtering along rows of an image followed by one-dimensional median filtering along the columns. Figure 40 illustrates this concept. Note that although

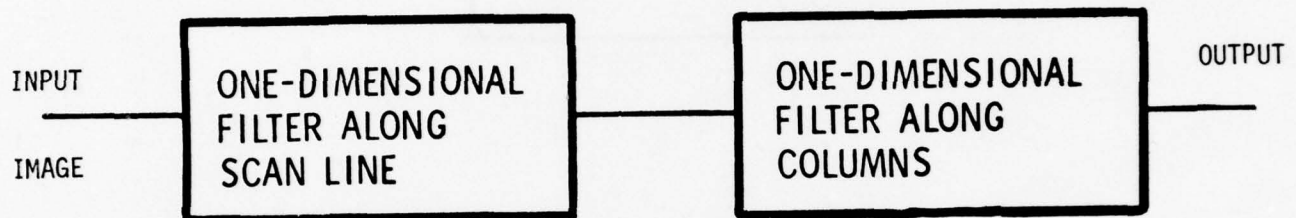


Figure 40. Separable Median Filter Concept

successive row and column filtering is implied, storing of the whole image between successive one-dimensional filtering is not necessary. This is shown in Figure 41 where two 5-point one-dimensional median filters replace a 25-point median filter. We should emphasize that, because of the non-linear nature of median filtering, the result of separable median filtering is not identical to the equivalent full two-dimensional median filter. But the advantages of the true two-dimensional median filter--namely spiky noise

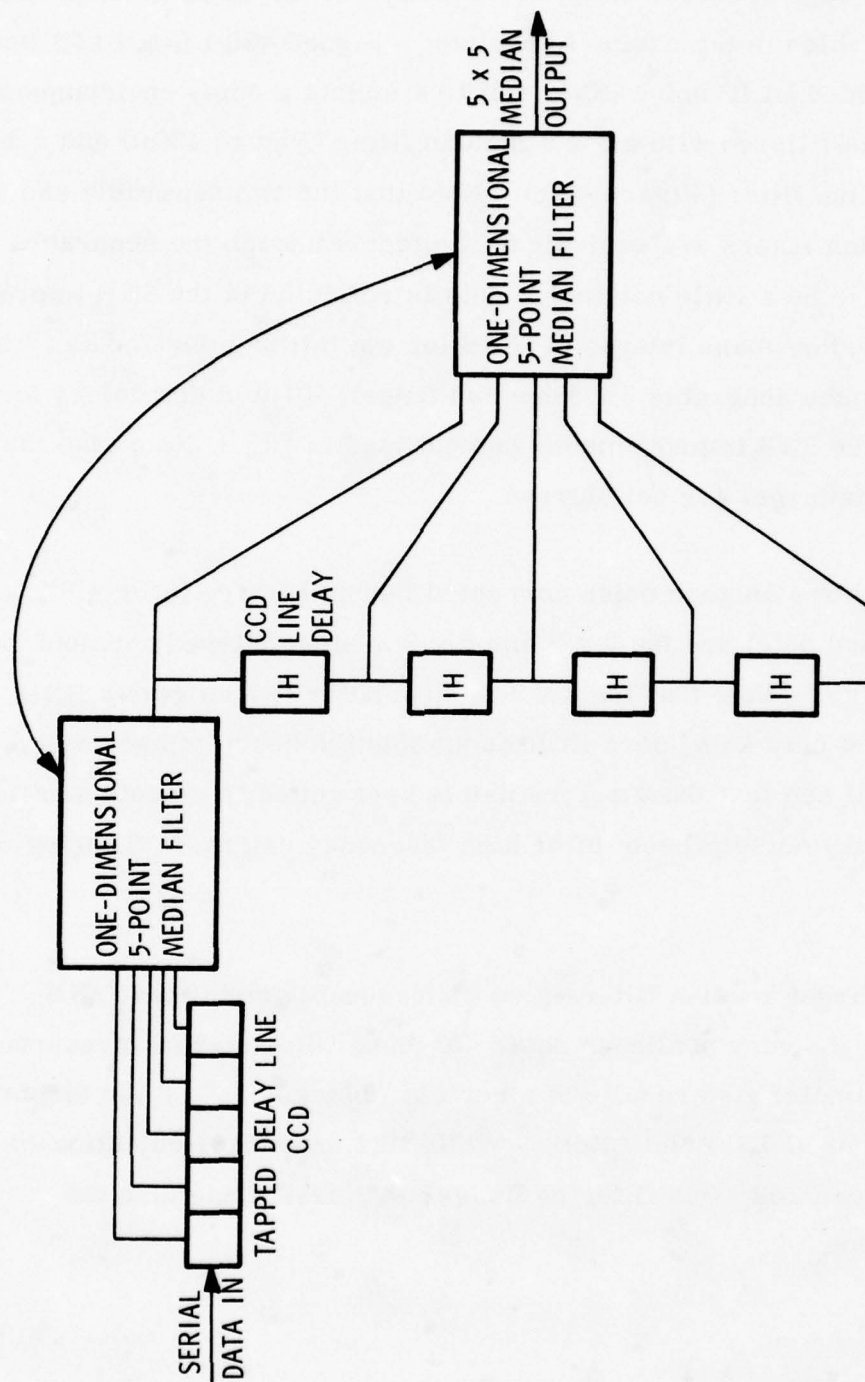


Figure 41. Separable Median Filter Structure

smoothing and edge preservation-- are equally preserved in the separable median filter which is far easier to realize. Figure 42(a) is a FLIR image to which we added BLIP noise ($\text{SNR} = 3$) to simulate a noisy environment. This image was filtered with a 5×5 median filter (Figure 42(b)) and a separable 5×5 median filter (Figure 42(c)). Note that the two separable and the full 5×5 median filters are equivalent in effect (although the separable filter appears to be a little noisier). This is reflected in the SNR improvements measured on these images (4.3 dB for the full 25 point median filter and 4.1 dB for the separable 5×5 median filter). (The methodology for measuring these SNR improvements was outlined in [2].) Note also that the edges in the target are not blurred.

Figure 43(a) shows another noise corrupted thermal image ($\text{SNR} = 3$), and Figure 43(b) and 43(c) are the 3×3 and 5×5 median filtered versions of this noise image. Note that the 3×3 median filter (which gave a SNR improvement of only 1 dB) does little to smooth the heavy image noise. Indeed, we will see that the 3×3 median is best suited to smooth small amounts of spiky noise present after high frequency emphasis filtering of FLIR images.

Although the larger median filters give a measurably significant SNR improvement, the very nonlinear nature of these filters (which preserves edges, for example) also results in a certain "blockiness" in the smoothed images (in Figures 42(b) and 42(c)). While this artifact is objectionable, it is not as predominant when filtering images with less noise than the examples shown.

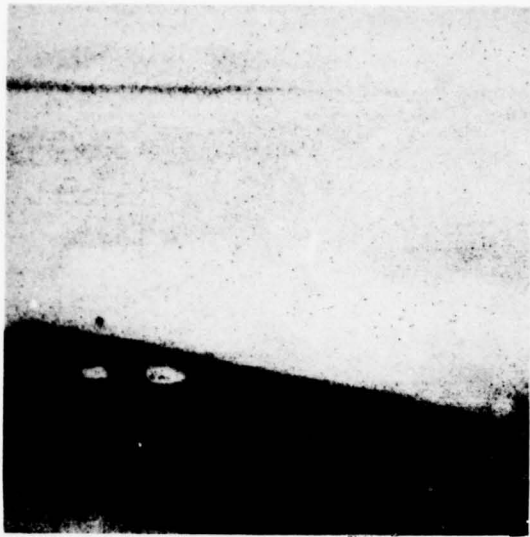


Figure 42(a). Thermal Image with Noise Added, $SNR \approx 3$

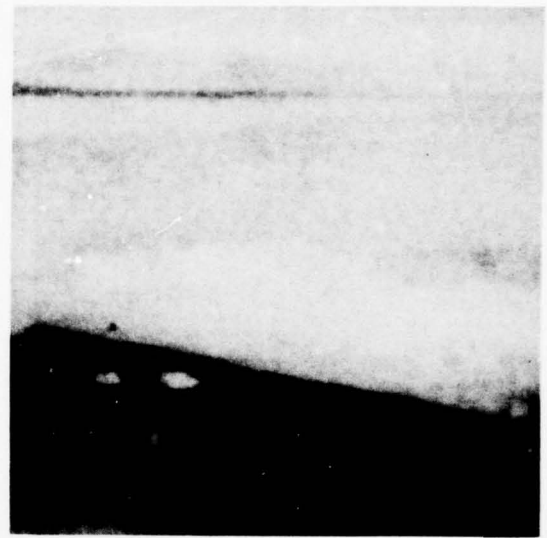


Figure 42(b). 5 x 5 Median Filtered

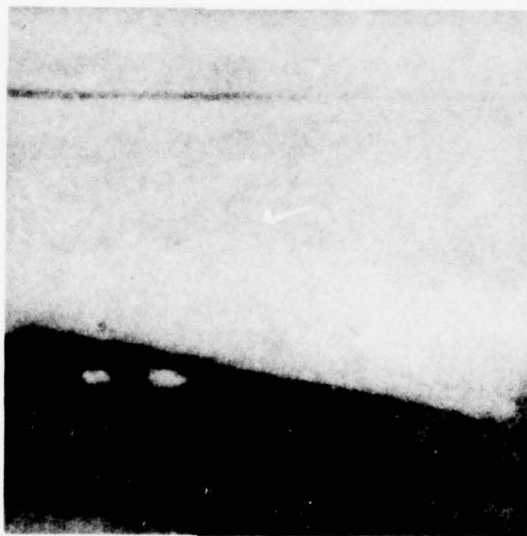


Figure 42(c). 5 x 5 Separable Median

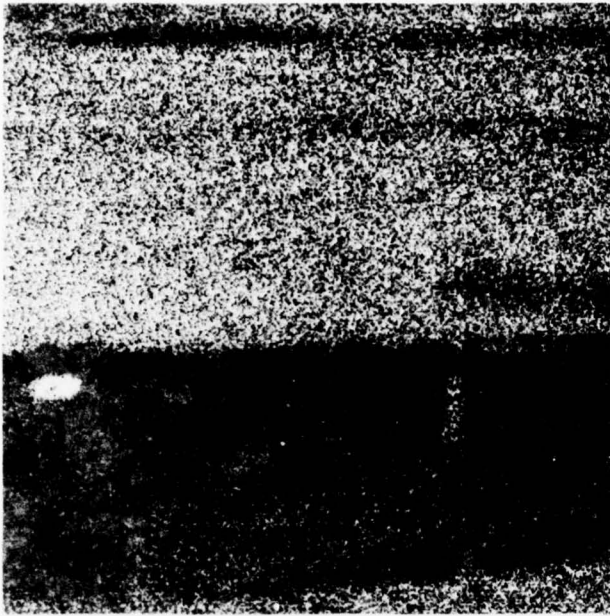


Figure 43(a). Thermal Image with
Noise Added SNR = 3

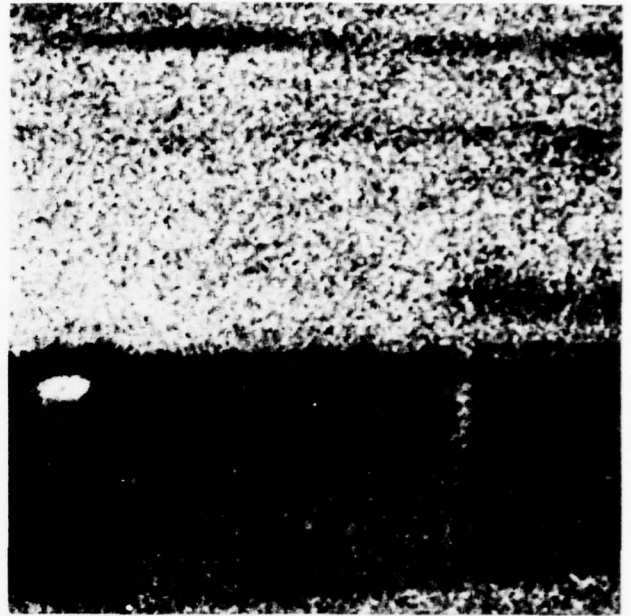


Figure 43(b). Median Filtered 3 x 3
S/N Gain 1.0 dB

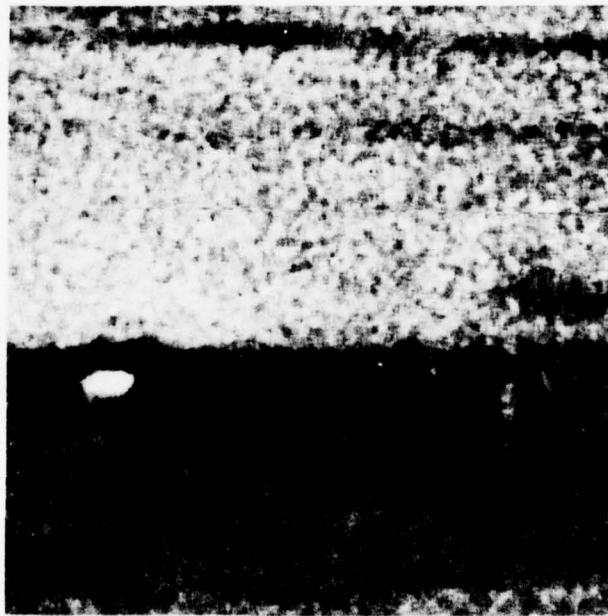


Figure 43(c). Median Filtered 5 x 5 S/N Gain = 4.3 dB

Adaptive Variable Width Filters

The problem of smoothing with a nonadaptive linear low pass filter is that while the overall SNR improves by a factor K (K is the size of an equivalent rectangular averaging window), the target edges become blurred. This is an undesirable effect, and this was one of the reasons why the median filter (which is a nonlinear filter) was investigated. An alternate approach is to use an adaptive linear filter that detects target and other significant edges and adaptively directs the degree of smoothing in each local area. We investigated adaptive variable width filters that are directed by: 1) the local curvature; and 2) the local gradient of the image.

The basic concept of these adaptive filters is shown in Figure 44, where a bank of different size filters is switched to realize the desired adaptive filtering. At each image point, the local gradient or curvature is computed in four principal directions. The maximum of these values then guides the selection of one of five smoothing filters (of increasing widths) to be applied at that point. For example, if there is an edge at point X oriented in any particular direction, the large value of the computed gradient or curvature at that point will cause a small window filter to be selected. Similarly, a larger window filter would be selected on an area with uncorrelated noise only, because this would give rise to a small local curvature or gradient. Thus, a scene directed adaptive filtering can smooth the overall image noise spatially without degrading the edges in the image. Two approaches to adaptive filtering were simulated that differed primarily in how the local image activity is measured; that is, the gradient directed and curvature directed adaptive filters. In the former, the local contrast determines the smoothing filter to be applied. In the second approach, the local curvature of the image intensity function directs the filter choice.

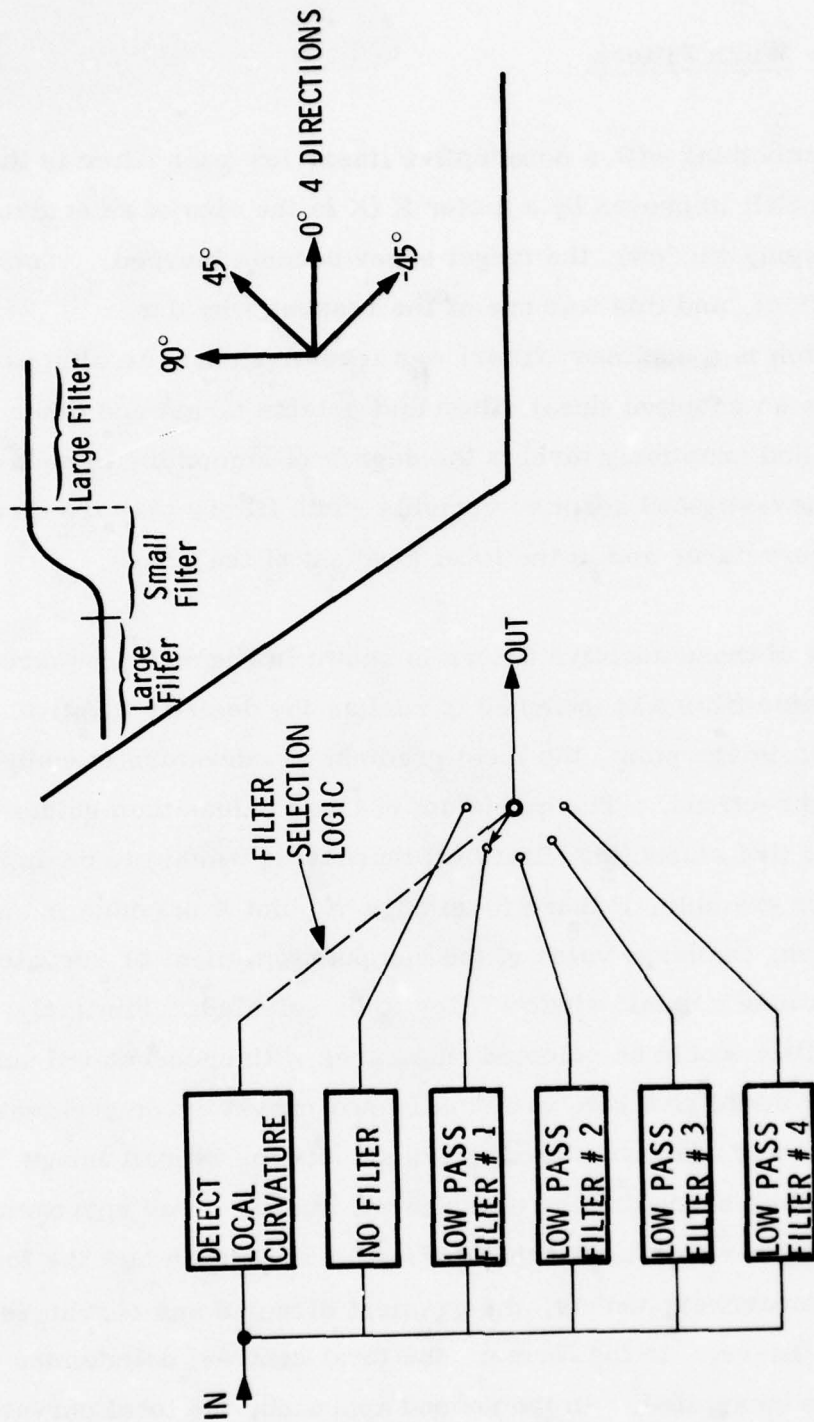


Figure 44. Adaptive Variable Width Intraframe Smoothing

Recursive and nonrecursive realizations of these two adaptive filters (curvatures and gradient directed) were simulated, thus resulting in an array of four different adaptive variable width filter algorithms. The details of these filter structures may be found in the interim report [2]. The nonrecursive filters here are similar in structure to the Gaussian low pass filters used in the section on contrast enhancement. The recursive filters are also similar, being first and second order two-dimensional Butterworth filters. The reasons for the alternate realizations remain the same; i.e., the recursive filters generally required fewer line delays and filter weights than their nonrecursive filters. Nonrecursive filter, on the other hand, can be isotropic, symmetric (zerophase) and in general are more commonly used in non-real time image enhancement.

The four adaptive filters were tested with several noise added thermal images to evaluate them and to select the best set of parameters for each filter. The improvement in SNR was measured by the mean square difference between the filtered image and the original noise free image. The details of this methodology may be found in the interim report [2].

The curvature directed adaptive filters were shown to be more attractive in [2] because the curvature (second order) is a truer measure of a local edge than the gradient (first order). In addition, the gradient and curvature directed filters performed roughly equivalently with the test images (interim report [2]). As a consequence, we selected the recursive and nonrecursive versions of the curvature directed adaptive filter for further evaluation. Figures 45(a) and (b) are the smoothed versions of the noisy image in Figure 41(a) filtered by the recursive and nonrecursive versions of the curvature directed filter, respectively. The gain in SNR for these filters

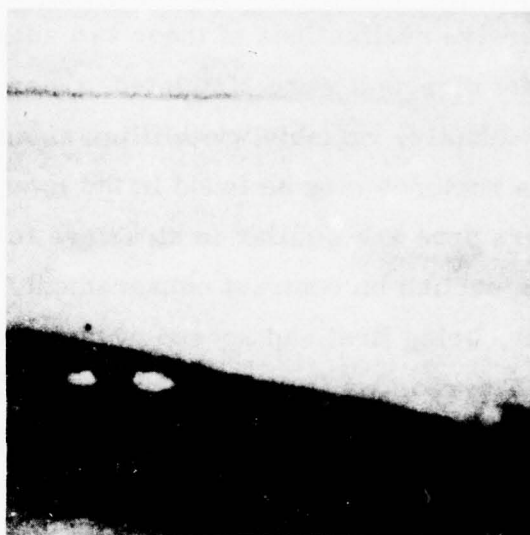


Figure 45(a). Recursive Adaptive Filtered Version of the Noisy Image in Figure 42(a) (S/N Gain = 2.13 dB)

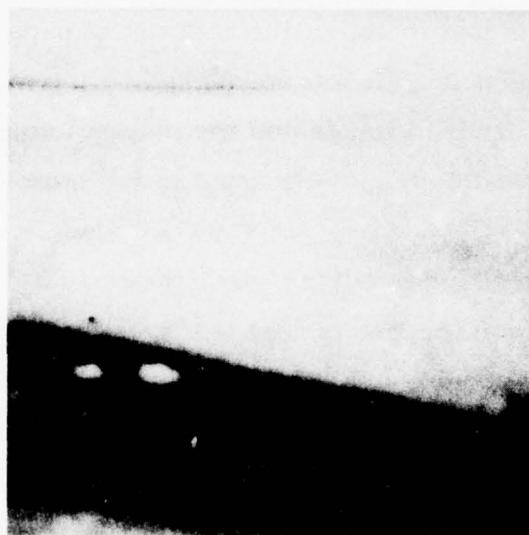
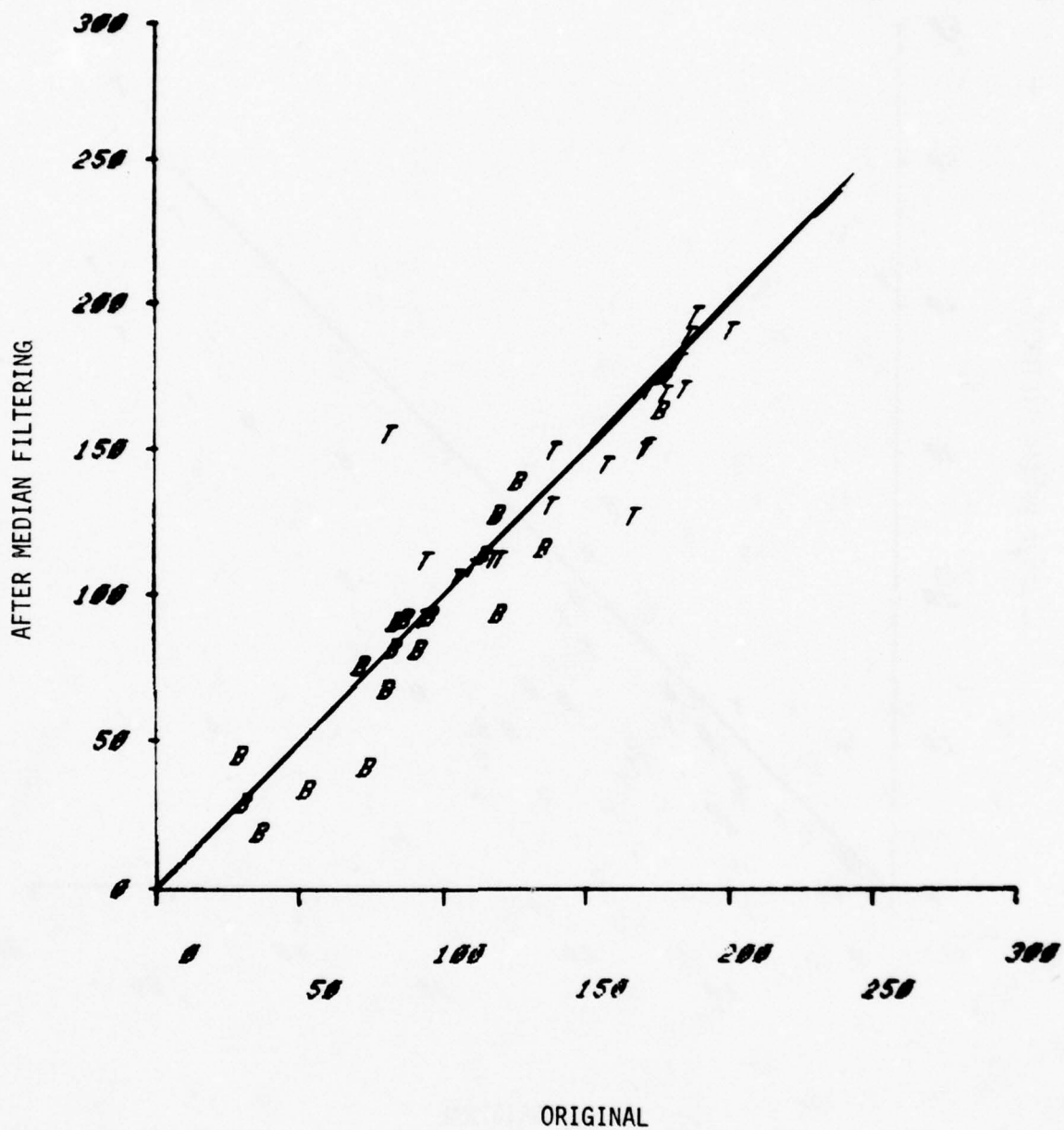


Figure 45(b). Nonrecursive Adaptive Filtered Version of the Noisy Image in Figure 42(a) (S/N Gain = 3.86 dB)



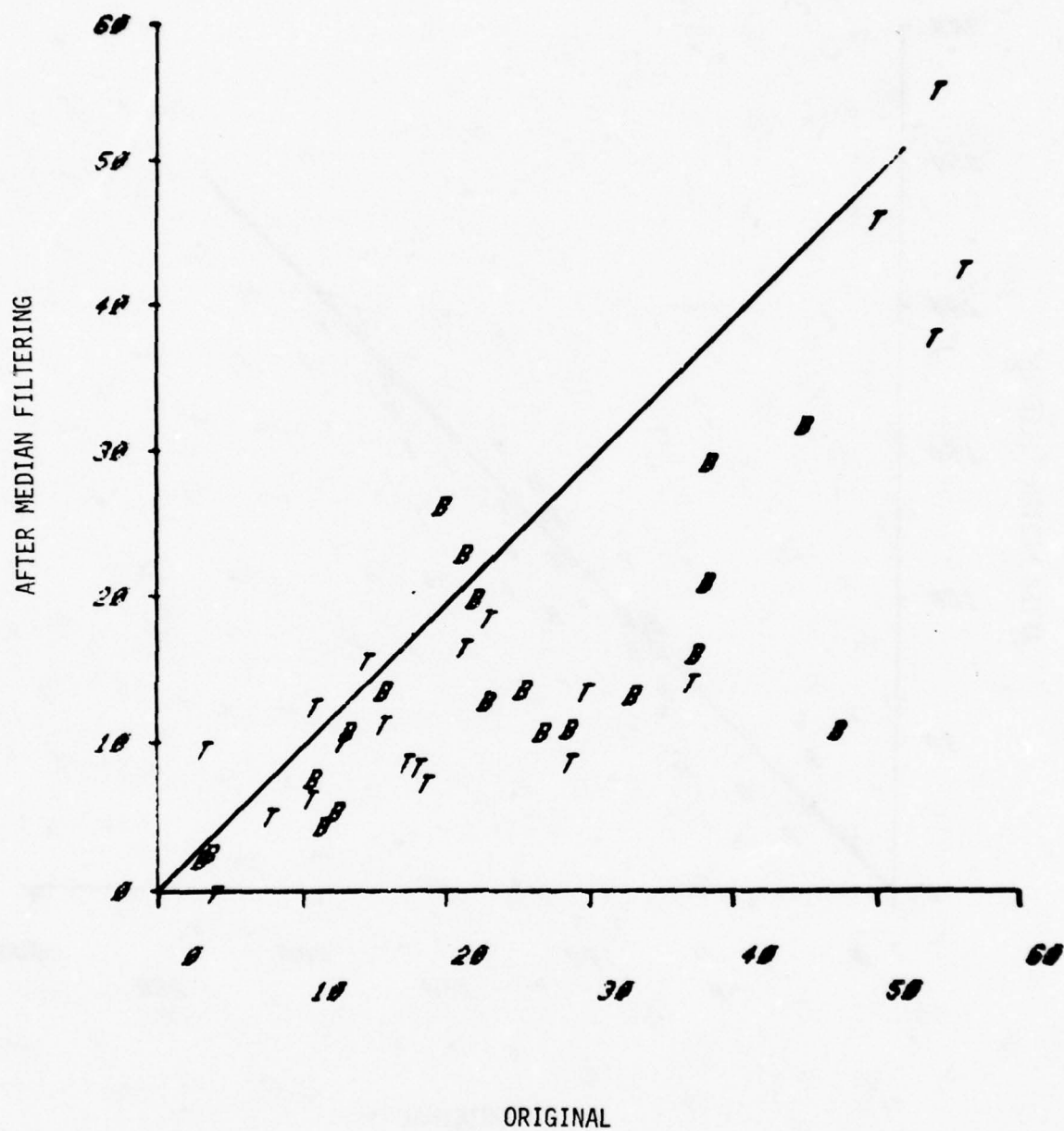


Figure 46(b). Standard Deviation of Intensity of Targets (T) and Backgrounds (B) Before and After 5 x 5 Median Filtering

is 2.13 dB and 3.86 dB respectively. Note that the nonrecursive filter smooths the image to a larger extent than the recursive filter. But the difference is due entirely to the difference in the chosen parameters for these algorithms and not due to the difference in filter structures (recursive and nonrecursive). Note that both adaptive filters have smoothed the subjective image noise somewhat better than the 5 x 5 median filter (although the measured S/N was approximately the same with the median filters).

STATISTICAL ANALYSIS OF INTRAFRAME SMOOTHING ALGORITHMS

As we pointed out above, we have an ensemble of several smoothing algorithms that includes the two-dimensional and the separable median filters, the hysteresis filter, and the four adaptive filters. Of these we eliminated the hysteresis filter as being ineffective, and we showed the separable median filter to be approximately equivalent in effectiveness to the full two-dimensional median filter. In addition, the gradient directed adaptive filters were considered inferior to the corresponding curvature directed filter. We therefore narrowed the number of intraframe smoothing filters for final evaluation to three:

1. Five-by-five median filter
2. Curvature directed nonrecursive (Gaussian)
3. Curvature directed recursive adaptive

These three filters were applied to the 40 thermoscope and FLIR images that constituted the test set. These filters were also applied to an additional set of 10 thermal images that were all corrupted with a known amount of white noise (SNR = 3). This was done to measure the amount of noise smoothing resulting from these intraframe smoothing filters. The resulting 10 noisy and 150 (50 x 3) filtered images were input to the human factors evaluation phase of the program.

In addition, we also measured the image statistics on these filters to quantify the effect of the noise smoothing algorithms. In particular, we can judge whether the contrast (especially edge and peak contrast) is adversely affected. In the same way, the effect on texture can be useful in determining whether local detail is blurred by these filters.

Intensity and Contrast Statistics

Figure 46(a) is the before versus after plot of the target and background inventory for the 5 x 5 median filter. The other adaptive filters exhibited similar behavior and they are not reproduced here. Note that the target and background cluster around the 45° line. This implies that the median filter does not significantly affect the mean target and background intensity. (Compare this with the contrast enhancement algorithms in Section IV). In Figure 46(b) we see the before and after plot of the target and background intensity standard deviation after median filtering. Here the points are below the 45° line, which means that the algorithms nearly always result in a lower standard deviation of target and background. This is to be expected, of course, from a noise smoothing algorithm.

Figure 47(a), (b) and (c) summarize the contrast features (peak, average, and edge contrast) measured on the three MRT enhancement algorithms. Comparing these with the corresponding curves for the contrast enhancement algorithms, we see that MRT enhancement algorithms do little to enhance the contrast measured. The median filter does better than the adaptive filters in the edge contrast measure. In general, however, the three filters do not differ drastically from one another.

Texture Features

Figures 48(a), (b) and (c) are plots of the "mean" texture feature (spacing $\Delta = 1, 2$ and 4) after median filtering versus the original texture measurements. Note that smoothing the image decreased the texture activity (as well as the noise we see for $\Delta = 1$).

INTRAFRAME MRT ENHANCEMENT SUMMARY

The most promising of the seven algorithms investigated were the 5×5 median filter, the curvature directed recursive adaptive filter and the curvature directed nonrecursive adaptive filter. The 3×3 median filter is too small to provide substantial image noise smoothing although it may be useful as a post filter after edge emphasis to remove the residual noise. The 5×5 median filter is difficult to implement in real time hardware; the equivalent separable 5×5 median filter is much easier to implement, however, and yields equivalent (although not identical) results. All three filters are roughly equivalent in the measured signal-to-noise improvement, although the median filters tend to produce a certain blockiness in the filtering noisy images. Therefore, the tradeoff will depend on the implementation

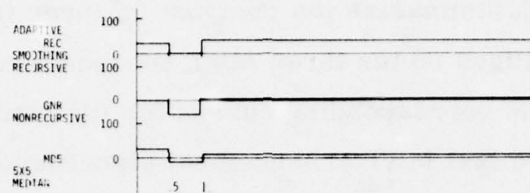


Figure 47(a). Percentage Change in Peak Contrast vs. Original Peak Contrast

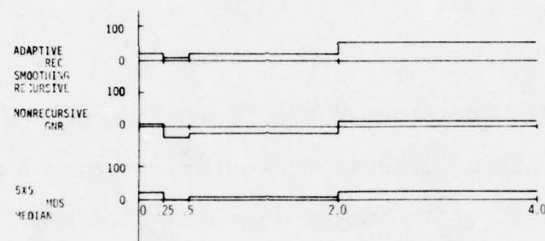


Figure 47(b). Percentage Change in Average Contrast vs. Original Average Contrast

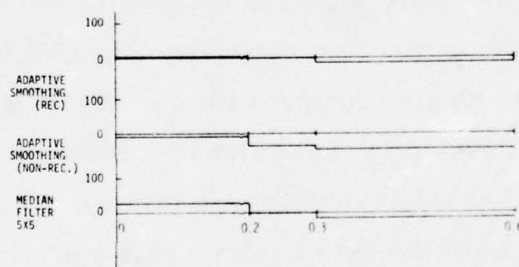


Figure 47(c). Percentage Change in Edge Contrast vs. Original Edge Contrast

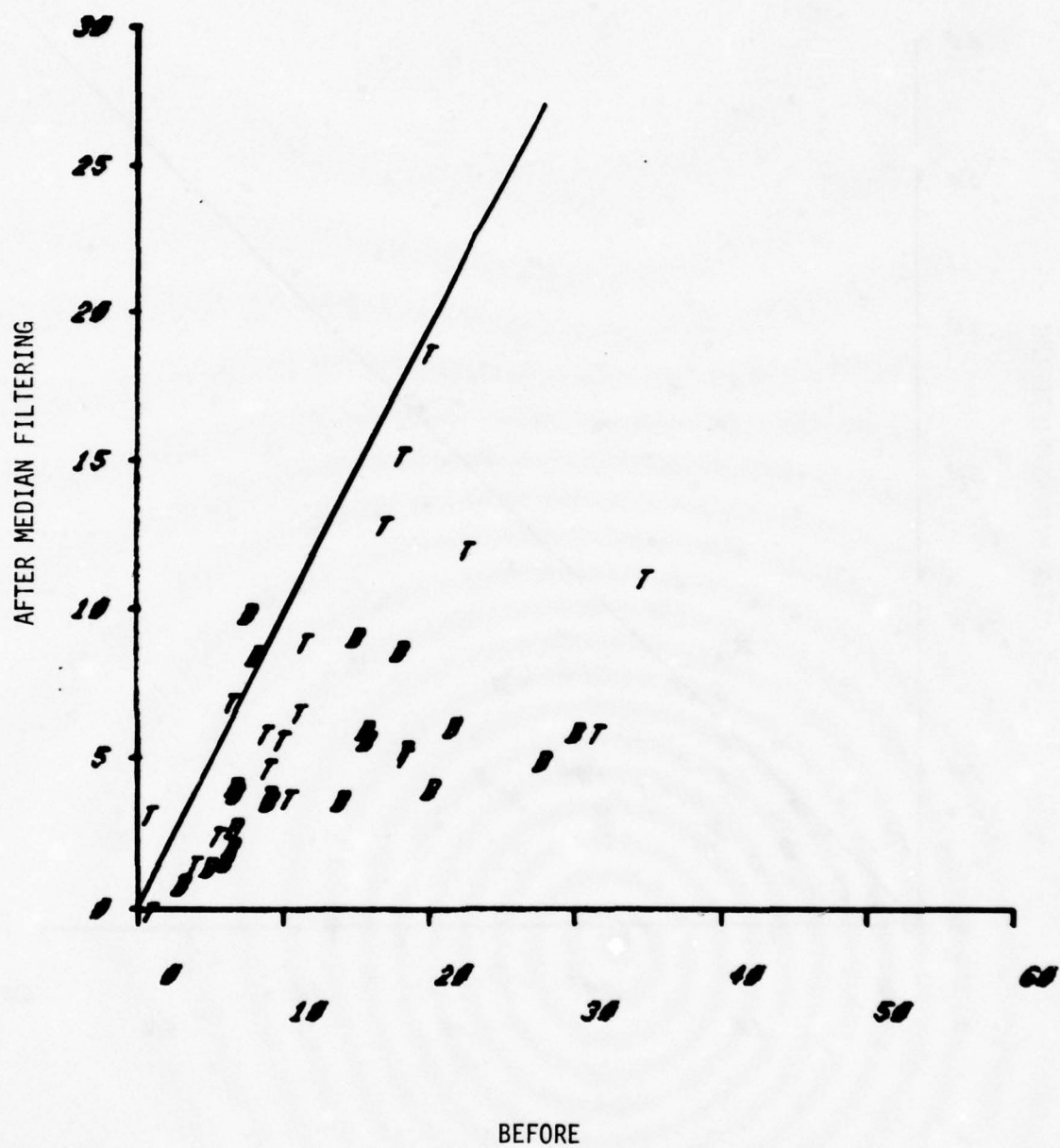


Figure 48(a). "Mean" Texture Feature. Before vs. After 5 x 5 Median Filtering (Spacing $\Delta = 1$)

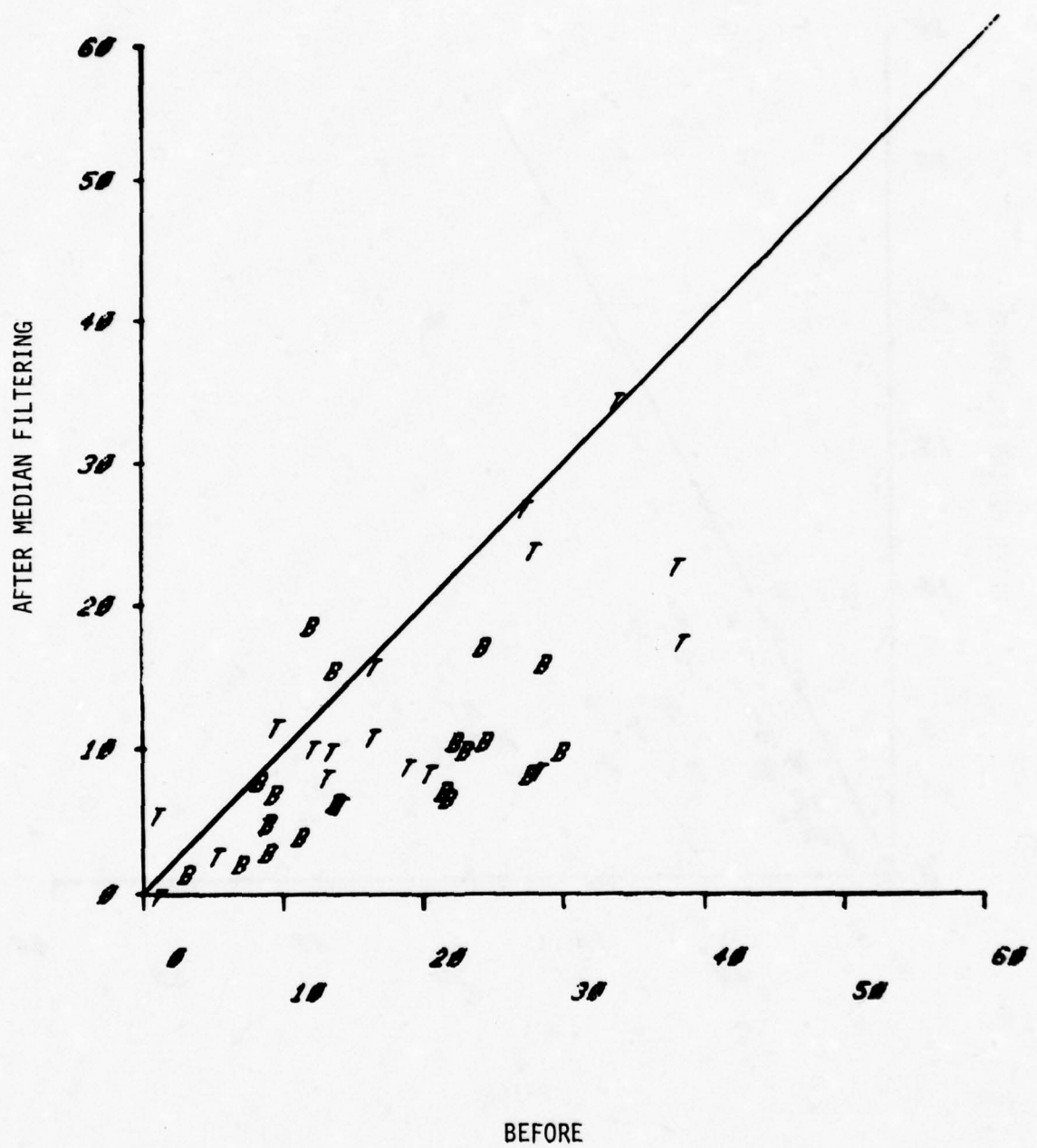


Figure 48(b). "Mean" Texture Feature. Before vs. After 5 x 5 Median Filtering (Spacing $\Delta = 2$)

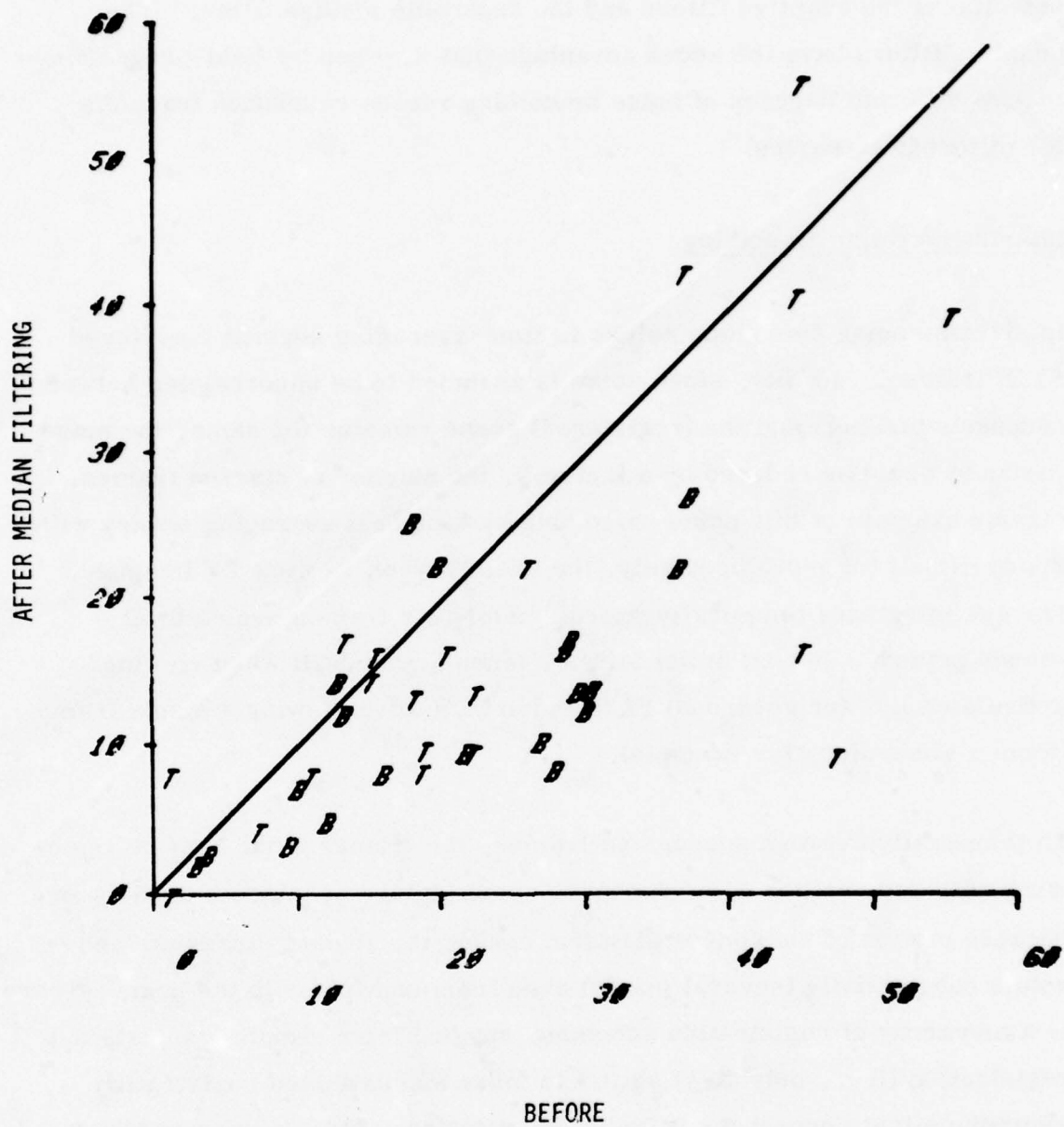


Figure 48(c). "Mean" Texture Feature. Before vs. After 5 x 5 Median Filtering (Spacing $\Delta = 4$)

potential of the adaptive filters and the separable median filter. The adaptive filters have the added advantage that they can be field-programmed to give different degrees of noise smoothing versus resolution tradeoffs for different scenarios.

Interframe Noise Smoothing

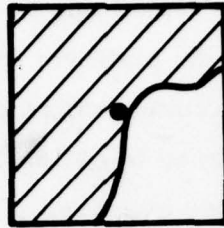
Interframe noise smoothing refers to time-averaging several registered FLIR frames. Ideally, since noise is assumed to be uncorrelated between successive frames and the (registered) scene remains the same, the noise variance would be reduced by a factor N , the number of stacked frames. A classic example of this noise smoothing by temporal averaging occurs with the eye itself (or more precisely, the brain), when viewing TV imagery. The eye integrates temporally approximately six frames over a 0.20 second period. Hence, it perceives a much higher SNR when viewing articulated live (or recorded) FLIR video than when viewing a single frame (from a video disc, for example).

To temporally average successive frames, the frames must first be registered with one another with respect to some feature of interest in the scene. because motion of the sensor platform causes the scene to translate and rotate substantially (several pixels) even from one frame to the next. There is a hierarchy of registration schemes, ranging from simple translational registration (i.e., only (x, y) shifts) to more sophisticated registration algorithms that correct for translation, rotation, scale and perspective transformations. Honeywell's Systems and Research Center and Radiation Center have been studying real time registration algorithms for interframe FLIR image enhancement. The details of these schemes and the simulation results can be found in the interim reports [2, 3]. We summarize these results here.

Translational Frame Registration--The simplest among registration schemes involves pure translational correction; i.e., no rotation or scale change. Such a scheme has been simulated by Honeywell with successive digitized and noise-added sequences of FLIR frames with tactical targets containing a stationary target (a truck) and another containing a target (an APC) moving at right angles to the platform trajectory. The registration was done using an automatic maximum correlation tracker algorithm against a corner of the target in each case, to estimate its translation from frame to frame. The frames were then averaged after being aligned to account for the target motion.

Figure 49 shows the correlation and reference windows used in the registration algorithms. The 5 x 5 reference window is chosen on the first frame in the sequence to encompass a high contrast target corner. Note that the 5 x 5 window can be at grid spacings of 1, 2, or 3, which means that the corresponding windows are actually 5 x 5 to 15 x 15 pixels wide in the image. The correlation window is a bigger window (11 x 11) in the next frame, chosen around the same center picture coordinates as the reference window. Both these windows are smoothed heavily with a 15 x 15 Gaussian filter of equivalent rectangular width of 7 x 7 prior to correlation. The correlation window is correlated with the reference window on Frame 2 and the translation (Δx , Δy) that yields the highest value of the correlation is found. This translation (Δx , Δy) is taken into account when shifting the second frame by this amount and adding to the first. The process continues with the point of maximum correlation on the second frame being the center of the reference window on that frame, and the third frame is correlated against this reference window, and so on. Two correlation criteria were

SMOOTHED
REFERENCE $r(p, q)$



5 x 5

Figure 49(a). Smoothed Reference Window

SMOOTHED
CORRELATION WINDOW $C_k(m, n)$

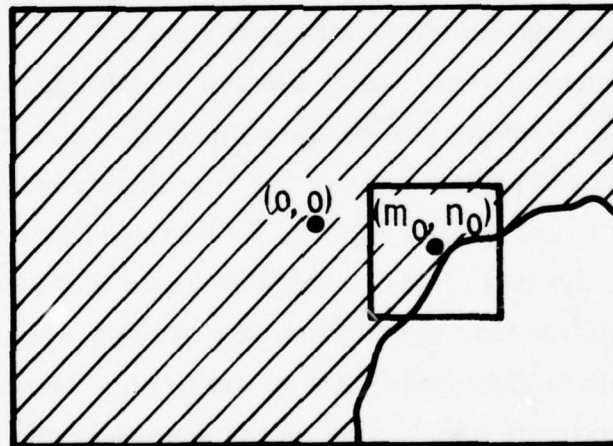


Figure 49(b). Smoothed Correlation Window Used in the Correlation Tracker Algorithm for Translational Registration

evaluated: the squared canonical product correlator (SCPC) and the minimum absolute difference correlator (MADC). The former has proved more robust in the presence of noise.

A Simulation Example--From a real time FLIR sequence, a set of nine noise-added (SNR = 3) FLIR frames 1/15 of a second apart were stacked and the S/N gain against the noise-free frame was measured at intervals of two, four, six and nine frames. The total integration time here is $9 \times 1/15 = 0.60$ second. Figure 50 shows the results of the interframe stacking for the stationary target case. Note that after stacking nine frames in Figure 50 the noise is reduced considerably (by 3 dB). Ideally, of course, when we stack N frames, we should realize gains in the SNR against the noisy image of $10 \log_{10} N$ dB because the noise is uncorrelated from frame to frame while the scene is assumed to be the same. In Figure 51, the solid curve shows the actual S/N gain measured (as previously described) over the whole frame versus the number of frames stacked, and the broken curve shows the theoretical expectation in the S/N gain. The reason we do not monotonically continue to realize increasing gains in SNR by stacking more frames (as predicted theoretically) is that the translational registration is not fully correcting the sensor platform motion because of skew and rotation. The attendant scene blurring causes the S/N gain measured to fall short of the theoretical.

Other simulations were also done to verify the effectiveness with 1) moving targets, 2) lower SNR conditions, and 3) registering frames that are further apart in time. The algorithms worked well under all of these conditions. The interim report [2] contains the details.



Figure 50(a). First Noisy Frame
SNR = 3



Figure 50(b). After Averaging Two
Noisy Frames



Figure 50(c). After Averaging Four
Noisy Frames

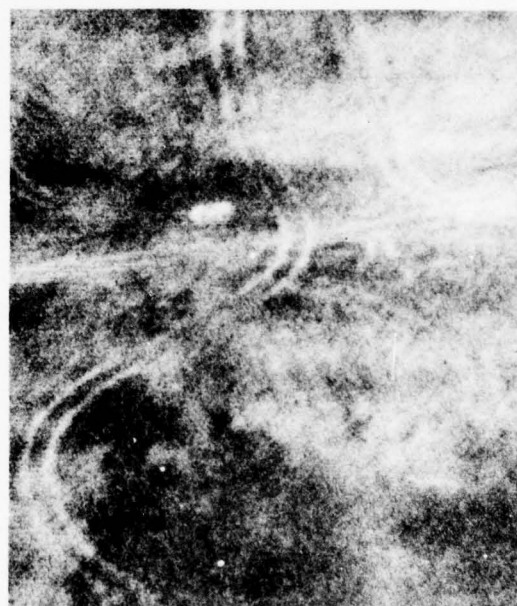


Figure 50(d). After Averaging Nine
Noisy Frames

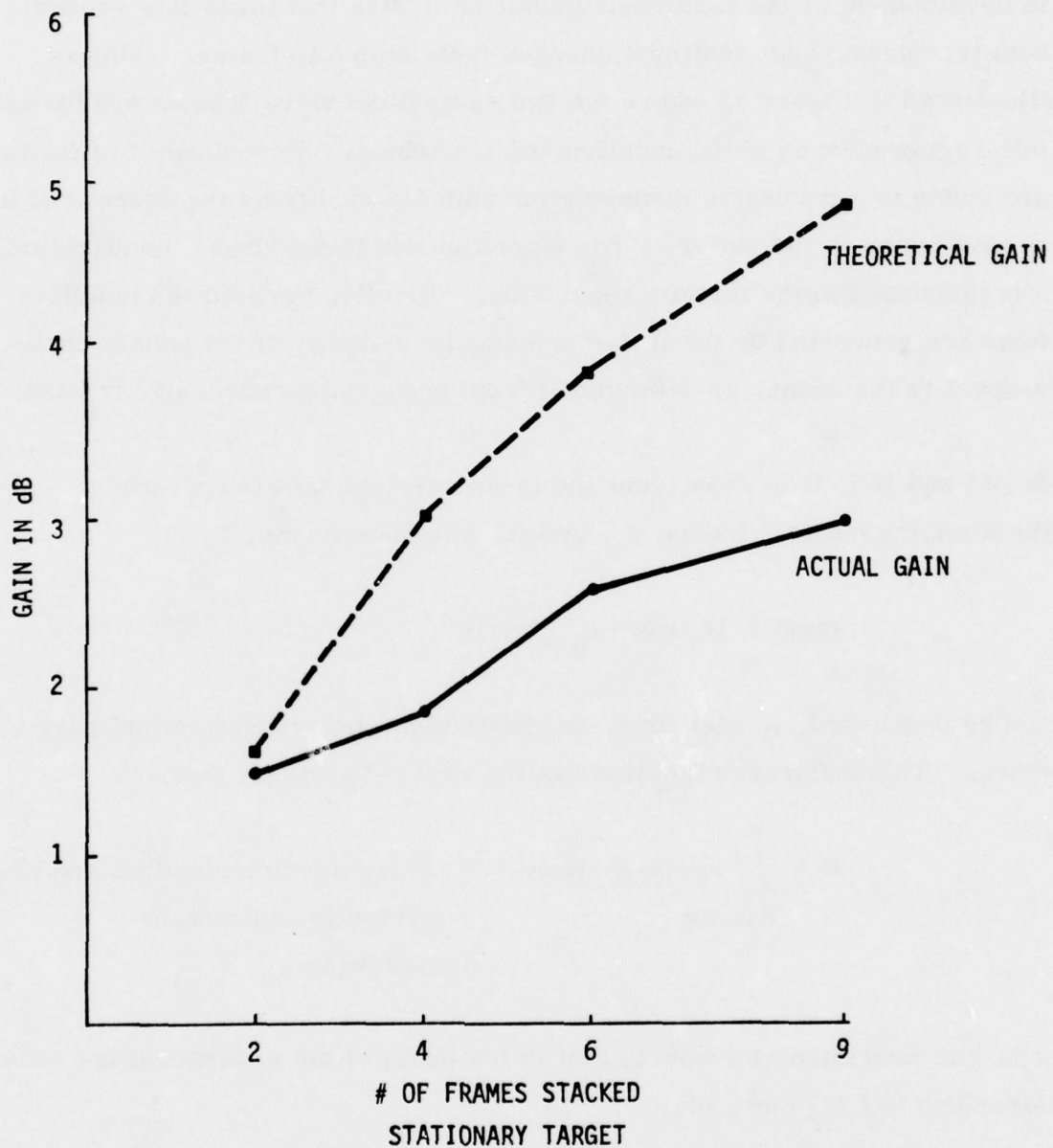


Figure 51. Signal-to-Noise Ratio Gains Accrued by Frame Averaging vs. the Number of Frames Stacked for the Stationary Target Case

Registration with Translation and Rotation--An alternate approach has been in development at the Honeywell Radiation Center that takes into account both translation and rotational changes from frame to frame. This is illustrated in Figure 52 where the two successive video frames are brought into registration by shift, rotation and stretching. Essentially two frames are added in a recursive manner after suitably modifying the address of the second frame in the buffer. The algorithm for this address modification was discussed in the interim report [3]. Briefly, the address modifications are generated by the linear and angular velocity of the sensor with respect to the scene, as determined from processing successive frames.

In [4] and [5], it is shown how the mean absolute difference (MAD) (between the current frame, f_n , and the previous frame, f_{n-1})

$$d(xy) = |f_n(xy) - f_{n-1}(xy)|$$

can be processed, in real time, to update eight distortion (motion) parameters. This difference function has the obvious property that:

$$D = \iint_{\text{Frame}} d(xy) \, dx \, dy = \begin{cases} 0 & \text{if frames are identical and are perfectly registered} \\ 0 & \text{otherwise} \end{cases}$$

It is this fundamental property that is the basis of the processing operations described in [29] and [30].

The squared difference algorithm (SAD) uses a similar function:

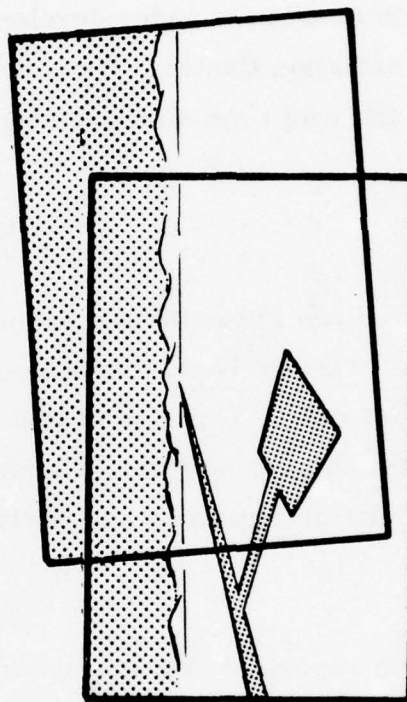
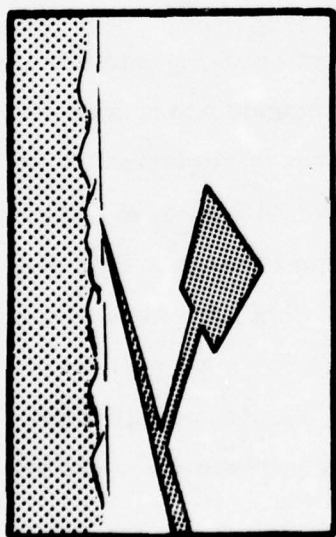
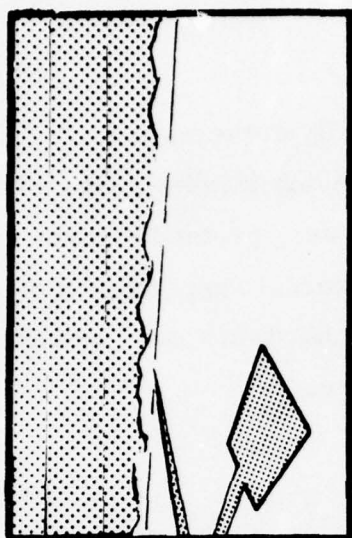


Figure 52. Two Successive Video Frames Brought Into Registration by Shift/Rotation/Stretching

$$d'(xy) = [f_n(xy) - f_{n-1}(xy)]^2.$$

This scheme for frame registration, which essentially corrects for platform motion only (and therefore cannot track moving targets), was not computer simulated under the current effort. However, prototype hardware for this function is under development (with internal support) at the Honeywell Radiation Center. When completed, this hardware can be used to evaluate the real time effectiveness of this scheme.

SUMMARY

We considered two alternate approaches to registering successive frames in real time. One evolved from simple tracker algorithms and corrects for translation only. This was simulated with noisy FLIR frames with moving and stationary targets. The second, which was designed to correct for platform rotation and scale change as well, was not computer simulated.

In the interim report [2] we also showed feasibility of tracking and stacking more than nine frames and frames more than 0.20 seconds apart with stationary and moving targets. The merit of this form of registration is that it is purely translational and can be accomplished with simple correlation type trackers. There already exist fully developed systems that can accomplish this tracking. The principal drawback of this approach is the following. The registration is done against a hot spot (or a target) by tracking its translation from one frame to the next. This means that the target has to be first acquired before tracking can be performed. However,

with the advent of the automatic target screener, this target acquisition function can be automated and even prioritized. In conjunction with the target screener, this form of translational registration should prove very tractable.

The question that comes to mind now is that since the human eye can track and temporally average six frames at 30 frames per second, what benefit does interframe averaging actually give us in a real time system? We proved feasibility of stacking frames more than 0.02 seconds apart (longer than the eye time constant) using the above tracking algorithms in the interim report [2]. Also, the registration techniques can be used with more than six frames (which is the limit of the human eye). The conclusion is that interframe registration and stacking would be a viable means of improving SNR in a noisy environment.

SECTION VI

RESOLUTION ENHANCEMENT

This section addresses two distinct problems relating to thermal image resolution: 1) The need to achieve full frame focus at all points in the field of view; and 2) the need to correct for the diffraction limited optics and detector blur.

The interim report [3] by the Honeywell Radiation Center extensively analyzed the focus problem in FLIRs and quantified the focusing difficulty encountered due to range and temperature effects. Range effect refers to the spread of the object ranges relative to the depth of field of the FLIR. It was shown that under certain circumstances these effects could be considered minimal. Temperature effects are due to thermal expansion or contraction of optical elements (particularly the germanium lenses used for thermal imaging). This creates focus problems in fixed focal length systems. Two schemes were described to correct this defect:

- Temperature compensation--the temperature of the objective lens is continuously sensed and this information is used to appropriately compensate the focus lens servo.
- Auto-collimation--a test image at infinity (or at some other specified range) is injected into the FLIR; the FLIR is then manually (or automatically) focused for this test image.

The following recommendations were made as a result of this analysis [3]:

For applications in which range-effects (on focusing difficulty) are minimal, consideration can be given to the use of either the temperature compensation, or of the auto-collimation, techniques. Both techniques have the potential for very simple implementation. The auto-collimation technique offers the additional operator benefit of a visual system-check feature. (If the special pattern at its usual clarity can be seen, the operator can be assured that the FLIR is operating correctly.) By eliminating the need for large-amplitude focus correction (to compensate for temperature effects on focus) the operator's focusing task is thereby reduced to only compensation for range effects. (The focus control should be calibrated in range; in many cases the operator need "dial-in" only a very approximate estimate of range to achieve adequate focus.) In this way the operator should be able to acquire a low contrast target directly without the wide-amplitude focus search procedures which can result in unreliable and slow acquisition.

For applications in which the focusing difficulty is due in large part to range effects (and suitable range information cannot be made available), then a true autofocus system is required. A combination of true autofocus with one of the other two techniques would improve focus acquisition (speed and reliability) if rapid (focus) acquisition proved to be a problem with low contrast imagery.

The mainstream of the image enhancement effort on this task was directed toward resolution restoration, i.e., correcting for the diffraction limited optics and detector blur over the entire field of view. Note that this assumes full frame focus has been achieved (either manually by step focus, or by autofocus) and the range effects are minimal, i.e., all parts of the image are in optimal focus.

The only remaining resolution degradation to be corrected is the diffraction limited optics and detector blur. This was addressed from an image processing point of view by modeling the optics and the detector transfer functions. Techniques to invert these degradations were explored.

The task of resolution restoration for thermal imagers can be divided into two areas:

- Full-frame resolution restoration for equalizing resolution out to the optical Rayleigh diffraction limit; and
- Superresolution to extend system resolution beyond the Rayleigh diffraction limit.

The reason for this dichotomy was that full frame resolution restoration up to the diffraction limit could be achieved by linear inverse filtering (in real time). Superresolution on the other hand, has always required iterative solutions--which means it cannot be real time. However, two promising near-real time solutions for superresolution were investigated: The gradient projection algorithm developed by Prof. Thomas S. Huang of Purdue University, and the Honeywell-developed stochastic approximation algorithm for super resolution.

OPTICS AND DETECTOR MODEL

The FLIR optics were first modeled to analyze their effect on the total system transfer function.

The space variant transfer function of the diffraction-limited optics was assumed (following common practice) to be Gaussian:

$$H(k_x, k_y | r_o) = \text{Exp} - \frac{k_x^2 + k_y^2}{2\sigma^2(r_o)}$$

where k_x and k_y are wave numbers.

Also,

$$\sigma(r_o) = \sigma_o + \sigma_2 r_o^2$$

where r_o is the radial distance from the optical axis, σ_2 was chosen so that the on-axis $\sigma_o = 4.2$, and the worst case diagonal $\sigma_{\max} = 3.54$. In the space domain, this corresponds to a worst off-axis blur wide diameter equal to 1.19 times the on-axis blur. This blur is obviously not very space variant, but it is representative of the actual ratios encountered in current FLIR optics and expected in future FLIRs.

Following the OTF, the detector induces a further blurring because the sensed output is the convolution of the image at the focal plane and the rectangular detector surface. The detector width was chosen so that it has the same area as the on-axis optics Gaussian blur (Figure 53). This corresponds to a detector width W given by

$$W = \frac{\sigma_o}{\sqrt{2\pi}}$$

This model essentially assumes that the detector induces approximately the same blur again as the optics. Figure 54(a) shows the combined OTF and detector MTF frequency response on axis and is seen to be very close to a Gaussian.

Referring to Figure 54(a), improving the system MTF up to f_0 can be done by inverse filtering. But beyond f_0 , we need to use the physical spatial constraints imposed on the restored image to resolve the image. Super-resolution purports to do this.

Note also that f_0 is the combined detector and optics cutoff. If we have an optics limited situation, the system f_0 is dominated by the optics diffraction limited cutoff. On the other hand, if the system is detector limited (i.e., the blur circle caused by the optics is small in comparison with the size of

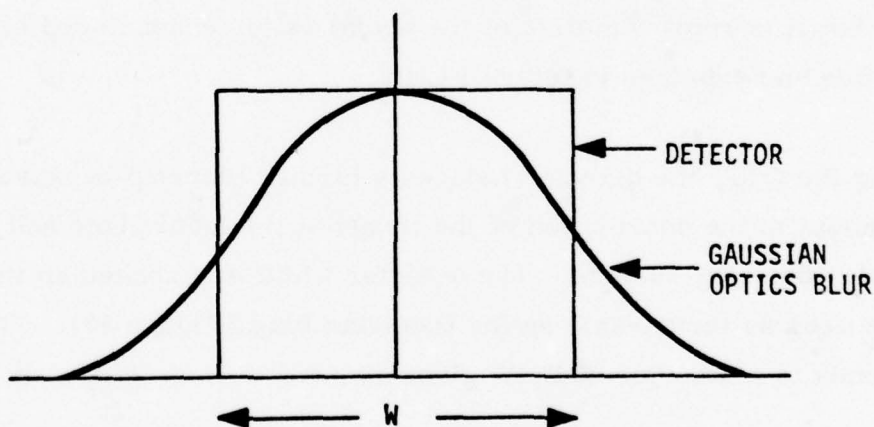


Figure 53. Detector and Optics Blur (Gaussian) Proportions Assumed.

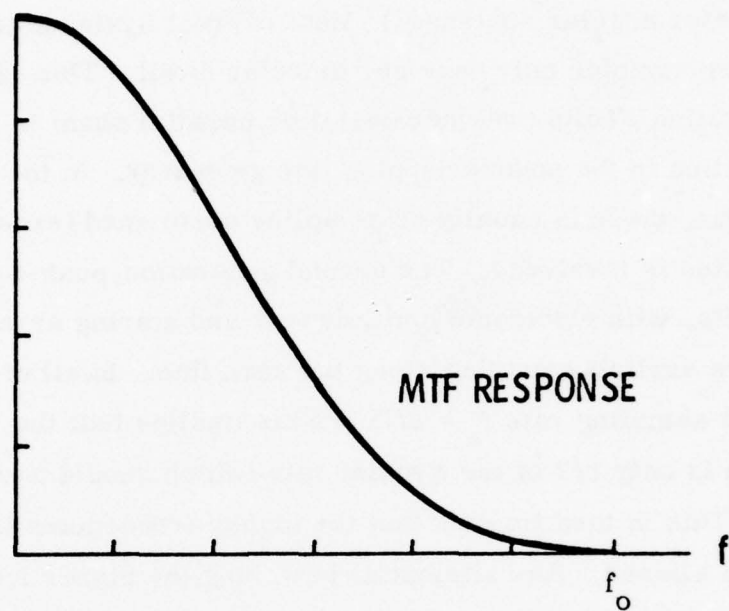


Figure 54(a). Combined Optics and Detector MTF

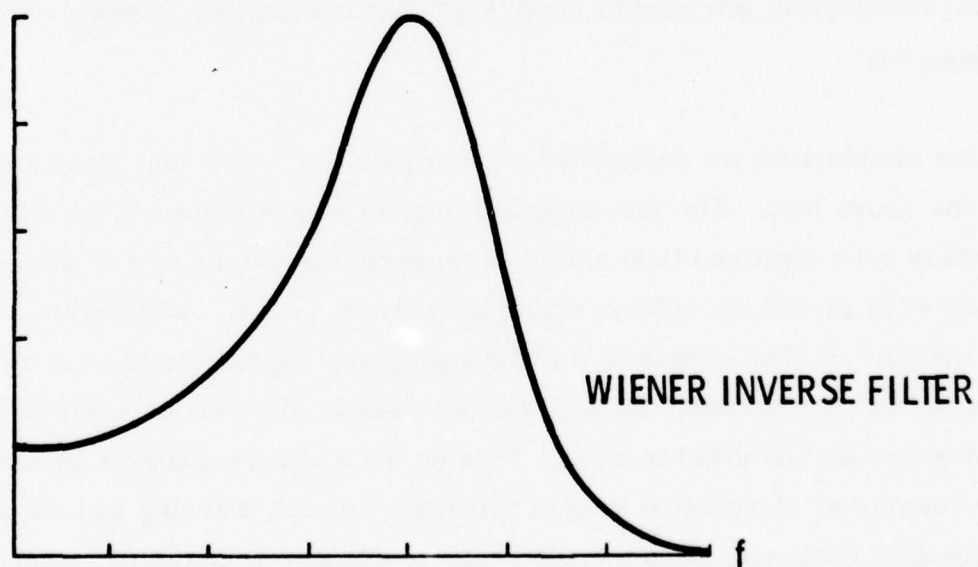


Figure 54(b). Wiener Inverse Filter Response

the detector), then the system cutoff f_0 is the detector sinc function zero $= \frac{1}{D}$. (D is the detector angular subtense.) Most current systems are detector limited and are sampled only once per detector dwell. This sampling in current generation FLIRs (whether serial or parallel scan) is implicit in the vertical direction in the nonoverlapping line geometry. In the horizontal (scan) direction, there is usually no sampling performed (except when time delay integration is involved). The second generation push-broom parallel scanning FLIRs, with electronic multiplexers and staring array configurations, perform explicit sampling along the scan line. In all of these systems, therefore, the sampling rate $f_s = 1/D$. This implies that the sampling rate is only $1/2$ of the Nyquist rate (which should have been $2 f_0 = 2/D$). This in turn implies that the higher frequencies in the sampled FLIR data are aliased. Any attempt at restoring the higher frequencies (by linear filtering, for example) will only accentuate these aliased frequencies. Therefore, there should be a caveat that when applying resolution enhancement techniques, one should make sure that the image has been sampled adequately.

In the simulations we performed on this task, we were constrained precisely by the above fact. The thermoscope images with which we were supplied had already been digitized (and sampled approximately once per detector dwell). They were therefore undersampled by a factor of two. In addition, the optics blur on this sensor (a thermoscope) was approximately the same size spatially as the detector size. In other words, the total blur due to the MTF was less than one pixel in area. This posed a severe obstacle in simulating the resolution restoration in both full frame linear filtering and the super-resolution schemes, because they rely on adequately sampled images. Consequently, with selected FLIR examples, we hypothesized further

blurring (e.g., five pixels) by the optics and detector MTFs. These degraded images represent an adequately sampled data set because the cutoff frequency of the degraded images is now less than one half the sampling frequency. They were then used in the simulations to evaluate the resolution restoration algorithms.

Inverse Filtering for Full Frame Resolution Restoration

Given an MTF as shown in Figure 54, we want to get back the attenuated higher spatial frequencies close to the system cutoff (first zero on the MTF). The simple inverse filter has the form

$$H'(f) = \frac{1}{H(f)}, H(f) \neq 0$$

where $H(f)$ is the degrading blur MTF of the system. The problem with the straight inverse filter is that it boosts the broad band noise at high frequencies where the signal power is essentially zero. Therefore, practically all linear restoring filters from Wiener on down have provision for making the inverse filter response small at frequencies where the signal energy is low. The Wiener (MSE) filter and Hunt's least square filter were studied on one-dimensional test patterns. For the simple analytic frequency limited blur MTF, they were found to be practically identical, and therefore we simulated the classical Wiener filter.

The details of this filter can be found in the interim report [2]. The inverse Wiener filter frequency response computed using a simple noise model is shown in Figure 54(b). We see that the Wiener filter is simply the inverse filter at frequencies where the object power spectrum is finite and nonzero,

and it tends to the blur response of system MTF at frequencies where the object power spectrum is null. Figure 55(a) is a reasonably sharp FLIR image that has been blurred with the linear space variant OTF with a large blur circle off-axis to model the space variant OTF. The on-axis blur circle was assumed to be five pixels in diameter. This image was restored by a piecewise linear Wiener filter similar to that shown in Figure 54(b) by dividing the image into sixteen segments, each being restored by one of four filters approximating the different blurs from on-axis to off-axis. Figure 55(b) is the resultant image, and we see that it has been restored adequately.

Although the above simulation assumed a space variant optics MTF, we noted before that this variation is not very significant. Therefore, we also simulated an inverse filter that was designed to invert the average blur over the entire frame and used the resultant filter over the full frame. The result was indistinguishable from Figure 55(b). A single space invariant filter should therefore prove adequate for full frame resolution restoration.

Summary and Conclusion of the Full Frame Resolution Restoration Task

We developed a methodology for linear inverse filtering of FLIR imagery when the blur is spatially variant. But the above study showed that the optics diffraction blur, for all practical purposes, can be assumed to be space invariant. Frequency domain inverse filtering is problematic for piecewise linear full frame focus restoration. This is because the narrow PSF translates into a broad MTF which requires a large area, two-dimensional discrete unitary transform (e.g., Fourier or Hadamard) to implement. The best approach may be to implement the inverse filter in a CCD analog spatial filter configuration (similar to the nonrecursive high frequency



Figure 55(a). FLIR Image Blurred with a Space Variant MTF Shown in Figure 22(a) (Five Pixels Blur on Axis and Seven Off Axis)

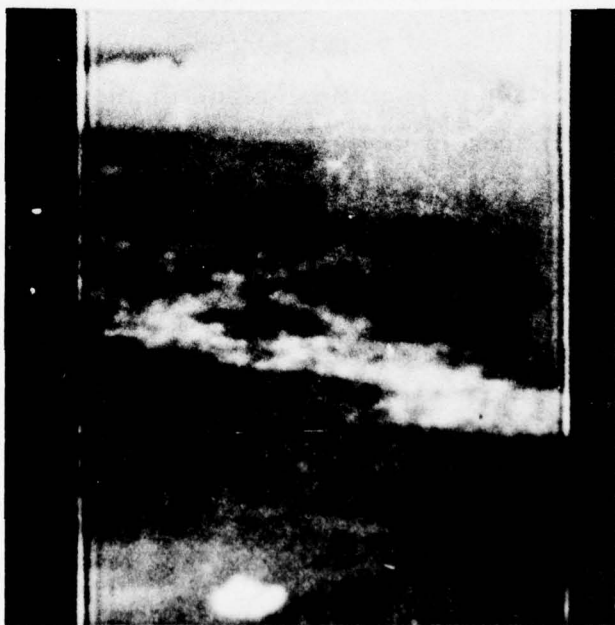


Figure 55(b). Restored by Piecewise Space Invariant Wiener Filter

emphasis filter discussed in the section on contrast enhancement). In addition, the space variant blur can be approximated by an average space invariant blur so that, for example, one analog filter can be used instead of four. We also note that the inverse Wiener filter shape in Figure 54(b) is very close to the high frequency emphasis filter in Figure 23(b). Therefore, simple high frequency emphasis could take the place of adaptive resolution enhancement filtering in a simple image enhancement design.

SUPERRESOLUTION ALGORITHMS

The linear, frequency domain resolution restoration techniques above can only restore the high frequencies out to the Rayleigh limit--the first zero on the OTF. Theoretically at least, extrapolation beyond the Rayleigh limit is possible due to the recovered information in the side lobes of the diffraction pattern. Several iterative techniques can use the positivity and boundedness constraints on the filtered output to accomplish this. These schemes are generally termed "superresolution" algorithms, and the most successful techniques for superresolution have been iterative processing algorithms.

Two iterative superresolution algorithms were investigated: the gradient projection (GP) algorithm (Huang, et al.) and the Honeywell-developed stochastic approximation (SA) algorithm. These algorithms were used to enhance NVL supplied thermoscope and FLIR imagery. The details may be found in the second interim report [2].

Digital Image Magnification

The iterative algorithms are non-real time and are computationally efficient. Their application therefore was envisaged in an essentially off-line mode and on a sub-image frame (e.g., containing an object of interest cued by the operator).

The section of the image containing the target in each image was magnified by a factor of two or three by linear interpolation, superresolved, and displayed in the magnified format. Magnification decreases the sampling grid size and so extends the sampling frequency. Two forms of the digital interpolation technique were simulated: 1) bilinear interpolation which is straightforward but has aliasing problems; and 2) a linear filter approach that overcame the aliasing problems using a separable Dolf-Chebyshev filter to perform the digital interpolation.

The magnified and interpolated sections were then superresolved using the GP and SA algorithms. The details of these algorithms are given in the interim report [2].

We found the SA and GP algorithms to be roughly equivalent in effectiveness (although the SA algorithm appears to perform slightly better); but the SA algorithm is easier to implement in a CCD processor because it requires one convolution versus two for the GP algorithm.

To apply the iterative superresolution algorithms to the magnified and interpolated thermal images, we first had to estimate the point spread function (PSF) of the original optics that degraded the FLIR images supplied

to us. Two PSF shapes, a rectangular PSF of width W and a triangular PSF of equivalent width W , were tested and found equivalent. In keeping with current generation FLIR optics, we also hypothesized blurs W of three, five and seven pixels on the magnified images. This corresponds to an optics blur of approximately 1.5 pixels (detector width) diameter before magnification. Note that this is an extremely small blur in relation to the sampling size. However, we applied the SA algorithm to several thermal (thermoscope and FLIR) images so that the results could be fed to the image enhancement evaluation task. (See Part II of this report.) These images were not further degraded before applying the SA algorithm in keeping with the philosophy of enhancing thermal imagery as acquired from current generation imagers without further degradation.

Figure 56(b) is a digitally magnified section of the FLIR scene in Figure 56(a) containing a 1.25 ton truck. This magnified image was super resolved using the SA algorithm and is reproduced in Figure 56. We had hypothesized the above blur function here, and we note that we gain only a slightly crisper version of the magnified image. This somewhat disappointing result is to be expected on two counts. We do not know the precise blur function of the images, and even if we did, the sampling rate of this image does not allow characterization of the blur. (The blur circle is about one pixel in diameter.)

Summary and Conclusions on Superresolution

As noted above, the results we obtained are not conclusive with respect to superresolution because the input images were not adequately sampled. However, the SA and GP algorithms proved effective when applied to simulated thermal imagery where the effective sampling rate was greater than the Nyquist rate.



Figure 56(a). Subframe of a FLIR Image Containing a 1 1/4 Ton Truck



Figure 56(b). Digitally Magnified
Version of Figure 56(a)



Figure 56(c). SA Algorithm Applied
to the Magnified Image
in Figure 56(b)

We cannot over-emphasize the importance of adequate sampling in thermal imagers if resolution enhancement is to have any impact on the design of second generation FLIRs. If FLIR designers insist on sampling once per detector dwell and future imagers continue to be detector limited as in the current generation, there is no hope for resolution enhancement. The only application we can see under these circumstances is when a very small aperture is dictated by size constraints. This could result in a system which is optics limited. Hopefully, in such a system, sampling once per detector dwell would be better than the Nyquist rate.

On the positive side, even in the absence of adequate sampling, real time digital image magnification itself could be a valuable aid to a FLIR operator. In addition, if the conditions for sampling are met in a future imager, the SA algorithm could be used to effect a near real time super resolution of a digitally magnified sub-image.

The scenario that envisages the use of superresolution in tactical imagers would be as follows. The operator cues a target area that is sampled at a higher rate and interpolated in two dimensions to obtain a finer grid to resolve the target. This area is then superresolved iteratively off-line and display magnified to the operator either on a split screen or separate display. Because of the amenability to CCD processing, these algorithms can be implemented in near-real time. The SA algorithm, for example, takes less than ten iterations to converge. At normal TV rates used for clocking the CCDs (30 cycles per second) this would mean one-third of a second--which is near enough real time to be useful. When augmented with automatic target cueing and digital image magnification, iterative resolution enhancement schemes would indeed be very useful for improving operator performance under tactical situations.

SECTION VII

INTEGRATED IMAGE ENHANCEMENT (CONTRAST, MRT, RESOLUTION)

In Sections 4, 5, and 6 we explored algorithms for contrast, MRT and resolution enhancement of FLIR images. These results were evaluated individually. However, in the image enhanced FLIR, these functions could conceivably be acting in tandem (e.g., in cascade). In addition, there was no guarantee that the cascade transfer functions could be optimal from an integrated system point of view. This provides the *raison d'être* for integrated image enhancement schemes that incorporate all three enhancement functions. This section consists of two logical subdivisions. The first summarizes the results of the cascade processing obtained by applying the above contrast, MRT and resolution enhancements algorithms back-to-back on FLIR imagery. The second is a development "from scratch" of integrated algorithms to perform all the enhancement functions. This latter work was performed by Prof. Thomas S. Huang of Purdue University under subcontract to the Honeywell Systems and Research Center.

CASCADE PROCESSING

After processing several FLIR frames (with the three contrast enhancement (CE) and three interframe MRT enhancement schemes, and the SA algorithm for resolution restoration), our next task was to combine one CE and one MRT enhancement algorithm to produce improved contrast and signal to noise ratio. These goals are unfortunately at odds with each other because, essentially, contrast enhancement algorithms boost higher frequencies,

whereas the noise smoothing schemes tend to suppress high frequencies. The MRT enhancement algorithms we have considered are adaptive, i.e., they were designed not to blur across edges and sharp details in the FLIR image while smoothing noise elsewhere in the image.

The order in which the CE and MRT enhancement algorithms are applied to a FLIR image is important. Should we smooth the image after contrast enhancing it, or vice versa? To put this question to rest, we subjected a thermal image to the following combination of algorithms and evaluated the results:

- 1) 3 x 3 median + local area gain/brightness control (LAGBC)
- 2) LAGBC + 3 x 3 median
- 3) 5 x 5 median + LAGBC
- 4) LAGBC + 5 x 5 median
- 5) Recursive adaptive smoothing filter + LAGBC
- 6) LAGBC + recursive adaptive smoothing filter

These results are reproduced in Figures 57(a), (b), (c), (d), (e), (f), (g). Figure 57(a) is the thermal image with LAGBC only. Note that some spiky noise is apparent in this contrast enhanced image. From the other images in this ensemble we note that the MRT filters applied before LAGBC result in certain contouring in the enhanced images. LAGBC followed by MRT

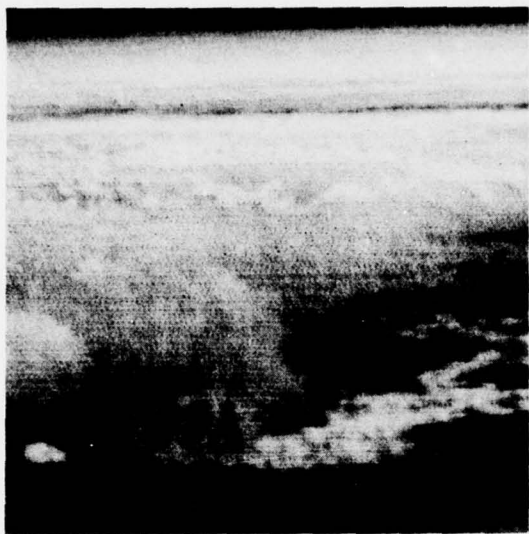


Figure 57(a). LAGBC Only

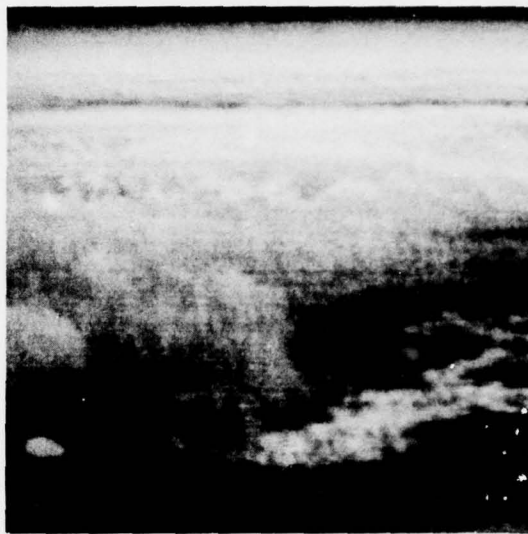


Figure 57(b). LAGBC + 3 x 3
Median

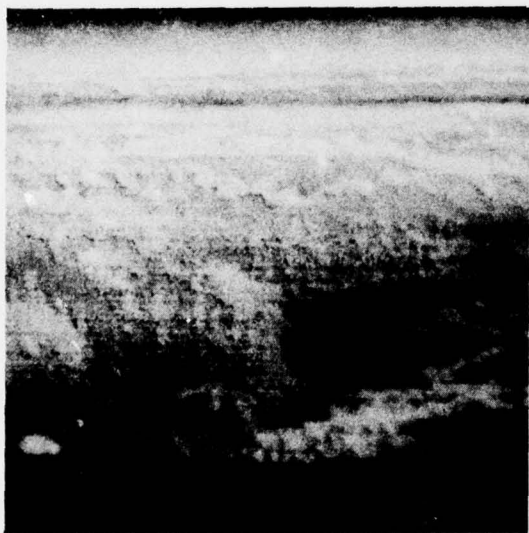


Figure 57(c). 3 x 3 Median +
LAGBC



Figure 57(d). LAGBC + 5 x 5
Median

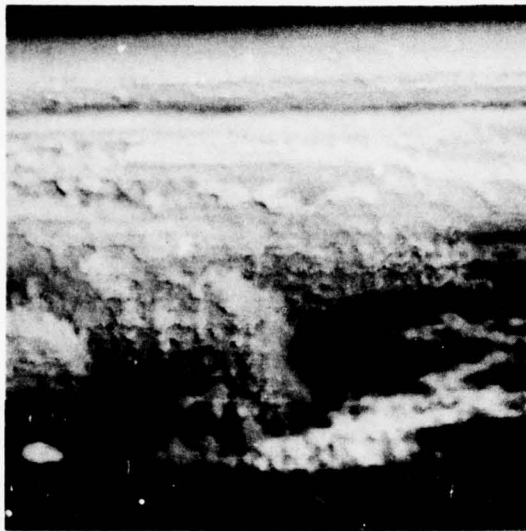


Figure 57(e). 5 x 5 Med + LAGBC

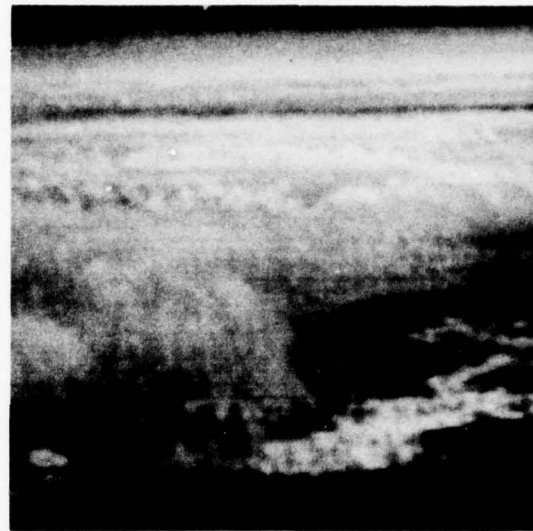


Figure 57(f). LAGBC + Recursive
Adaptive Smoothing
Filter

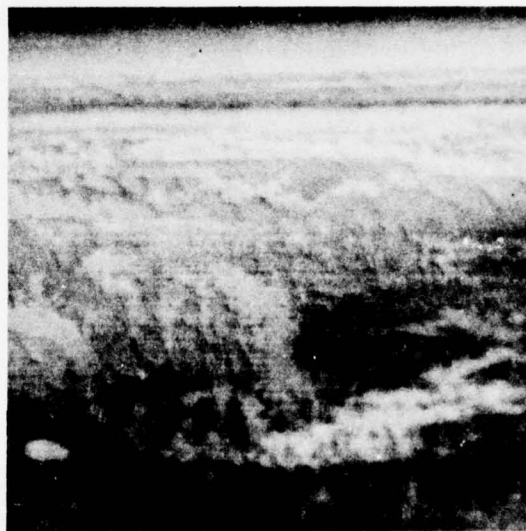


Figure 57(g). Recursive Adaptive Smoothing Filter + LAGBC

(smoothing) results in much more pleasing imagery. We also found that the 3 x 3 median filter has blurred the LAGBC image the least.

The conclusion from this test was that contrast enhancement should precede smoothing in cascade enhancements. In addition, the 3 x 3 median filter affords convenient spike noise removal after LAGBC without blurring the high frequency detail as would a larger filter. Note that the LAGBC followed by recursive adaptive smoothing filter fared extremely well. It also performed much better than did the 3 x 3 median filter (as we saw in Section V) in smoothing very noisy images. The combination of LAGBC + recursive adaptive smoothing filter was therefore used in the cascade production runs, which enhanced 20 images in this manner as input to the image enhancement evaluation process. The flowchart in Figure 58 shows the details of this cascade production run. In order to give the noise smoothing (MRT) algorithms a fair trial, the 10 thermal images input to these runs were complemented with the 10 noise corrupted versions of the same images. These 10 noisy images were also evaluated with the three MRT enhancement algorithms individually (5 x 5 median, recursive adaptive, and nonrecursive adaptive filters). These results were summarized in Section V.

We see in the flowchart in Figure 58 that digital magnification and the SA resolution restoration algorithm follow the other two enhancement algorithms. Because the resolution restoration algorithm is iterative, non-real time, and displays only a small portion of the enhanced image, this is the only logical place for these processes in the cascade, i.e., just before display. We see in Figure 57 that we get as natural byproducts of this cascade trial (at no extra cost): CE only, CE + MRT, CE + MRT + magnify images, and

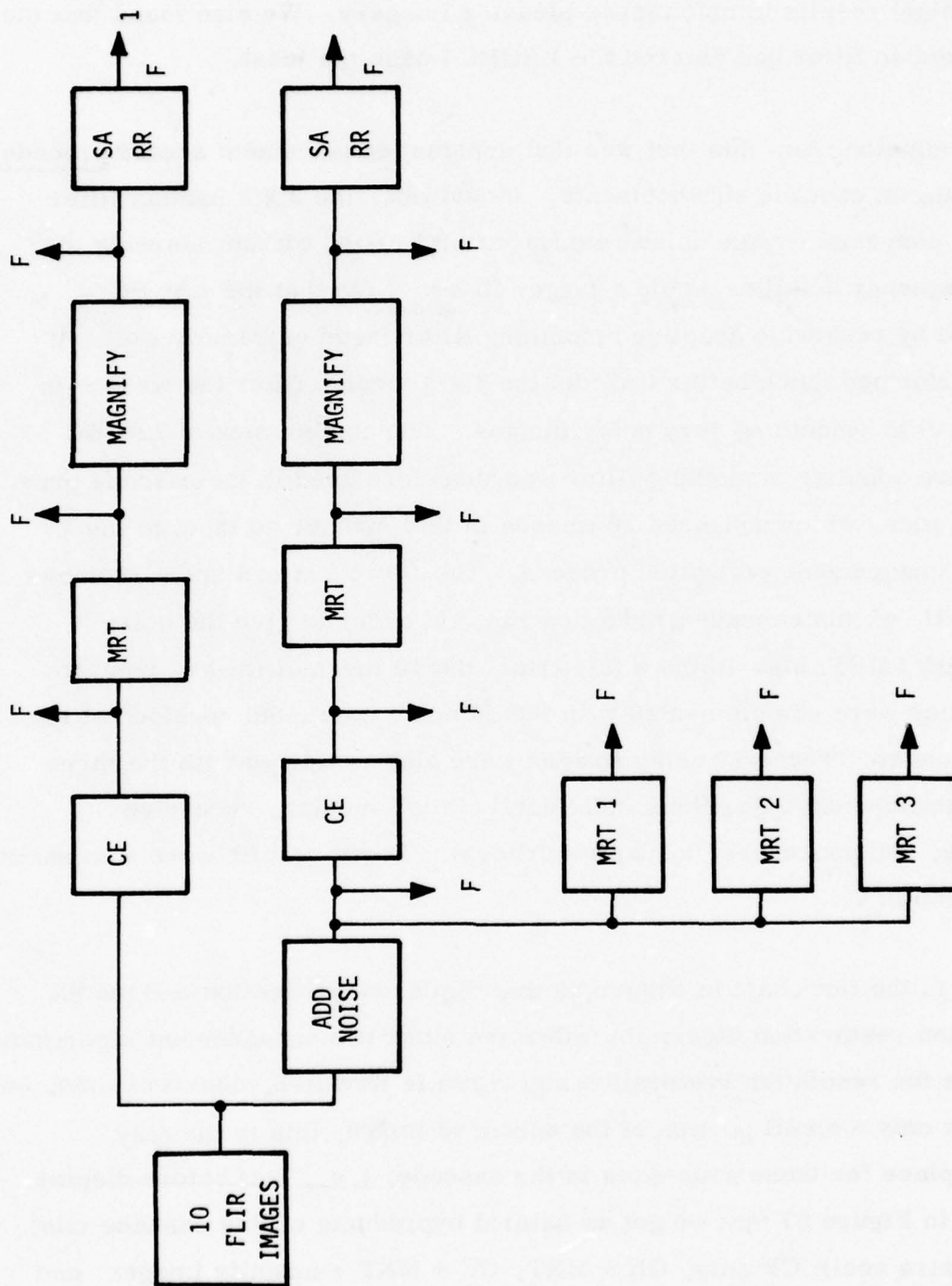


Figure 58. Flow Chart of Cascade Processing for Human Factors Evaluation

noisy, noise + CE, noise + CE + MRT, etc. A total of 140 images resulted from this trial that were subsequently evaluated by the human factors evaluation study. (Part II of this report).

An example of the result of this cascade trial is shown in the sequence of images in Figure 59(a), (b), (c), and (d). These are the noisy, noise + LAGBC + MRT and noisy + LAGBC + MRT magnified, and the corresponding resolution restored images, respectively. We observe that the cascade processing has indeed improved the noisy image. Note that while resolution restoration has done little to improve the magnified image, the magnified image has improved over the previous stage.

INTEGRATED APPROACH

Prof. Thomas S. Huang, under subcontract to Honeywell Systems and Research Center, explored an integrated approach to FLIR image enhancement. His approach differs from the cascaded approach above in that he is not confined to cascading individual algorithms. This scheme is parallel in nature, as opposed to sequential cascade processing. We include Professor Huang's results here.

Two-Channel Image Enhancement Algorithms (by Thomas S. Huang)

In a previous report* we described an integrated approach to image enhancement. The general algorithm contains two parallel channels. One channel

*Thomas S. Huang, "An Integrated Approach to Image Enhancement," in Honeywell's Interim progress report [2].

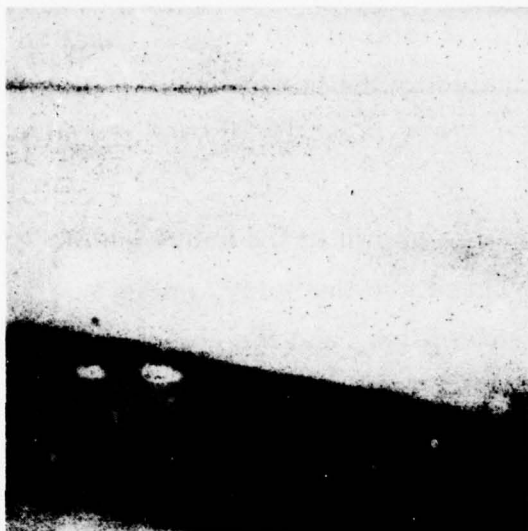


Figure 59(a). Noisy Thermal Image

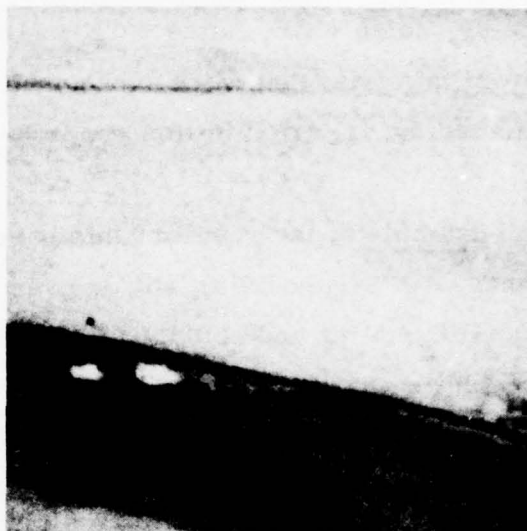


Figure 59(b). LAGBC + MRT



Figure 59(c). Digital Magnification



Figure 59(d). SA Resolution
Restoration

calculates a low passed version of the input. The other channel extracts edge information from the input and uses that to improve the quality of the output of the first channel. In this report we present the results of applying two special cases of this general algorithm to four FLIR test images.

Algorithm 1--A block diagram of Algorithm 1 is shown in Figure 60. To reduce computation time, very simple filters are used for H_1 and H_2 . The low pass filter is an equal-weight average over a 3×3 window in the spatial domain. The band pass filter H_2 is the difference between an equal-weight 3×3 average and an equal weight 11×11 average. Noise reduction in the output of H_2 (which is bipolar) is accomplished by band thresholding: the output at any time instant is set to zero if its value lies in the band $(-\theta, +\theta)$; otherwise it remains unchanged. The final output $\hat{f}(x, y)$ is a weighted sum of the outputs of the two channels.

Algorithm 2--A block diagram of Algorithm 2 is shown in Figure 61. The gradient-directed adaptive filter is the same one as described in Section 4 [1]. An edge-directed 3×3 median filter is applied to the output of the adaptive filter to get the final output $\hat{f}(x, y)$. Guided by the edge detector of the lower channel, the median filter processes only those picture elements which are around the edges in the image; thus it smooths the edges while leaving the rest of the image alone.

Results--Four FLIR test images are used. Each contains 512×512 picture elements with eight bits per picture elements (grey-level: 0-255). The two algorithms described in Sections II and III are programmed on a PDP 11/70 computer and applied to the four images. The results are displayed on a

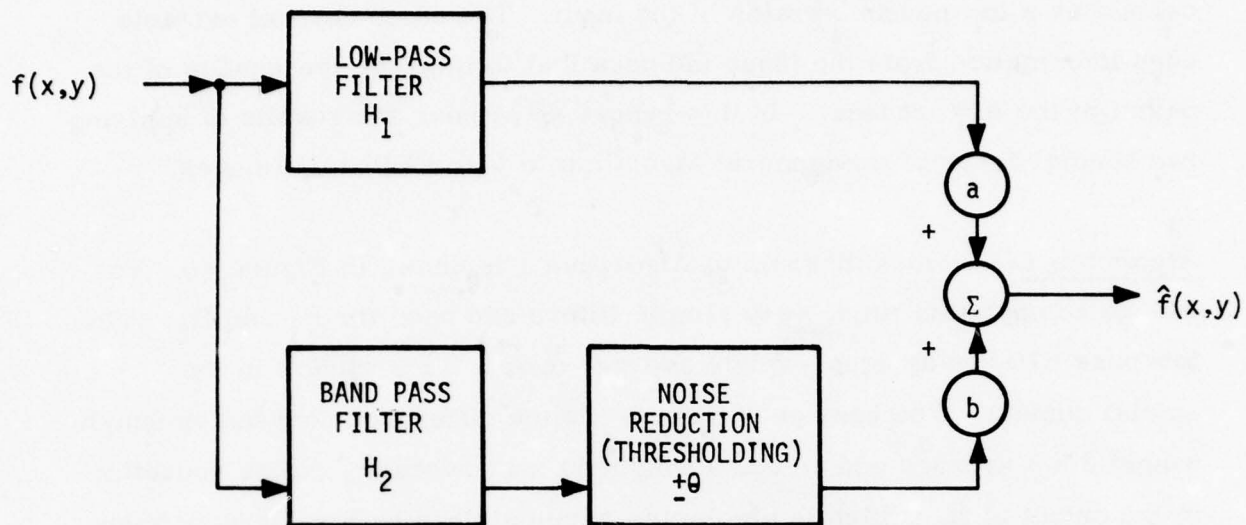


Figure 60. Enhancement Algorithm 1

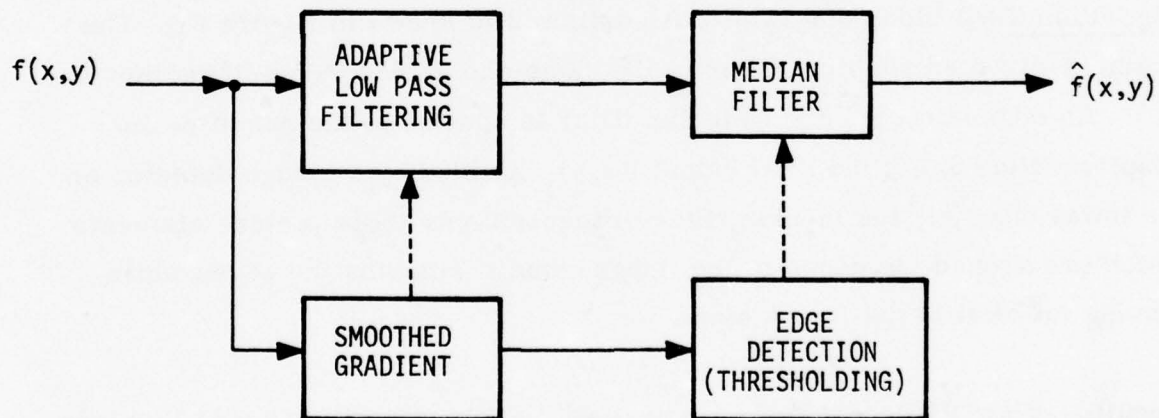


Figure 61. Enhancement Algorithm 2

RAMTEK and photographs taken therefrom. In all cases, the values of the parameters in Algorithm 1 are set to:

$$\theta = 15, a = 1, b = .75$$

We examine the processing of one of the images in detail. Figure 62 illustrates the application of Algorithm 2. Figure 62(a) is the original. The output of the adaptive filter is shown in Figure 62(b). Figure 62(c) shows the edge regions where the median filter is operative. Finally, Figure 62(d) shows the output of the median filter.

The application of Algorithm 1 to the same image is shown in Figure 63. The output of the 3 x 3 low pass filter is shown in Figure 63(a). The output of the band pass filter before and after thresholding are shown in Figures 63(b) and (c), respectively. Figure 63(d) shows the final output.

Omitting the intermediate steps, we show the results on two other images in Figures 64 and 65, respectively. In all cases, (a) is the original, (b) is the output from Algorithm 2, and (c) is the output from Algorithm 1.

SUMMARY AND CONCLUSIONS

We saw that cascading CE and MRT algorithms yielded a combination that enhanced the contrast and smoothed the image noise without interfering with each other's function. This is primarily because of the adaptive nature of these algorithms. We discovered that, for best results, CE should precede adaptive noise smoothing. The median filters proved valuable for filtering out spiky noise present in contrast enhanced thermal images. The combination selected for evaluation, however, was LAGBC + recursive adaptive



Figure 62(a). Original FLIR Image



Figure 62(b). Output of Adaptive Filter

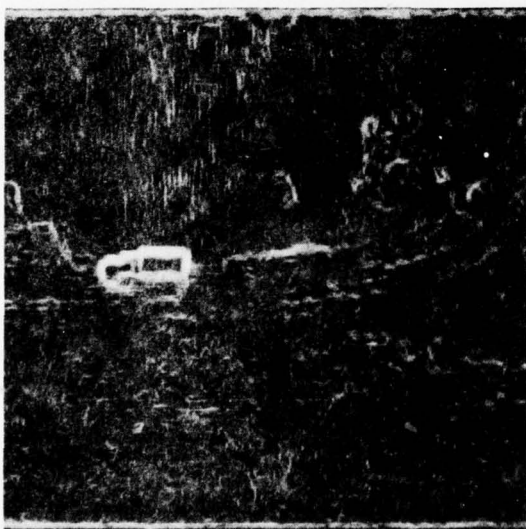


Figure 62(c). Edge Regions Where the Median Filter is Operative

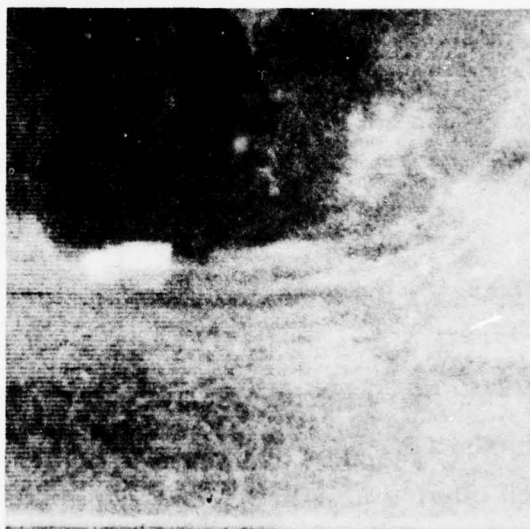


Figure 62(d). Final Output of Algorithm 2

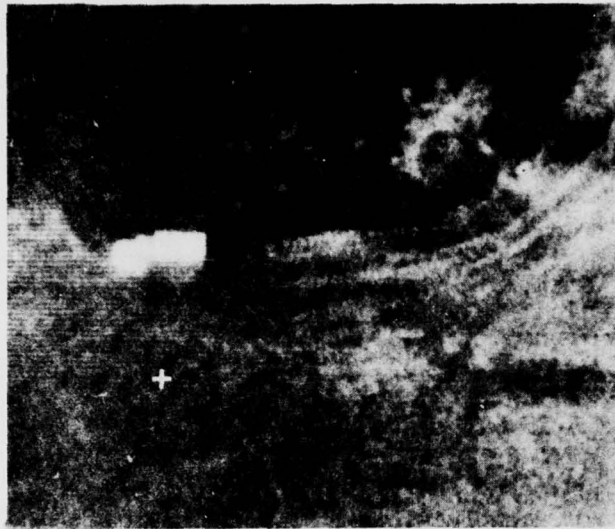


Figure 63(a). Output of 3 x 3 Low Pass Filter

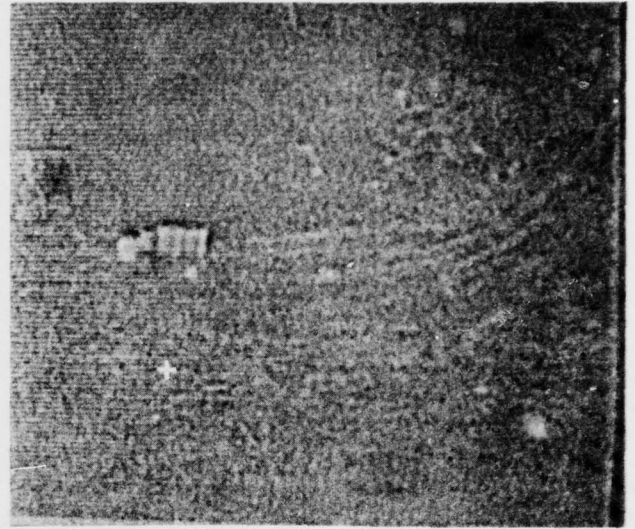


Figure 63(b). Band Pass Filter Before Thresholding

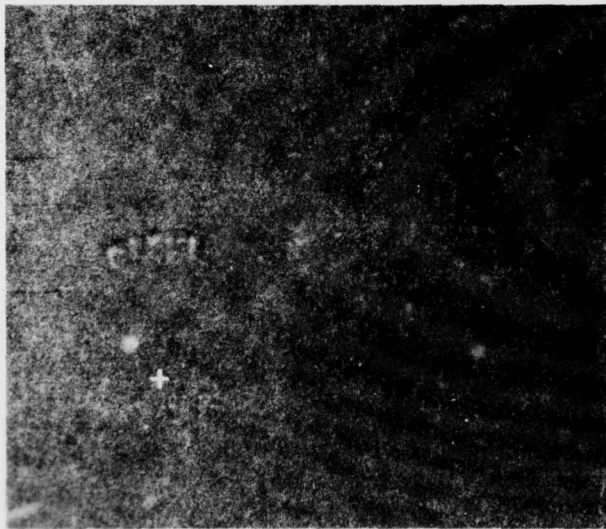


Figure 63(c). Band Pass Filter After Thresholding

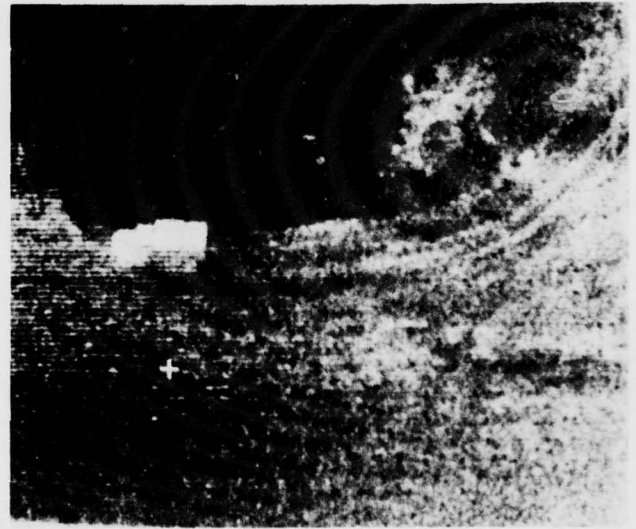


Figure 63(d). Final Output (Algorithm 1)

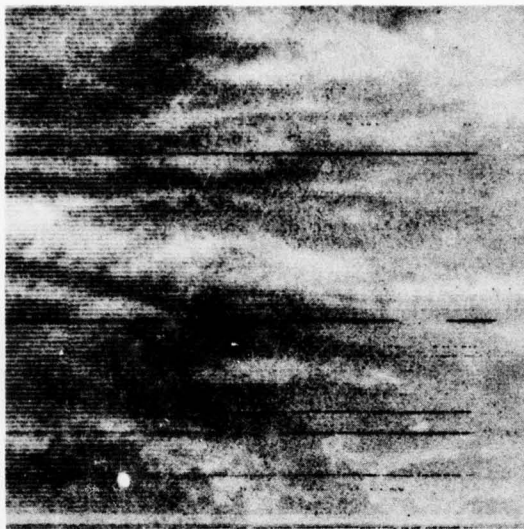


Figure 64(a). Original



Figure 64(b). Integrated Algorithm 2

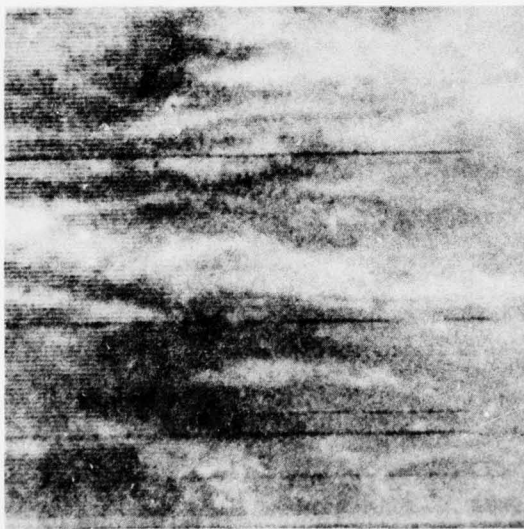


Figure 64(c). Integrated Algorithm 1

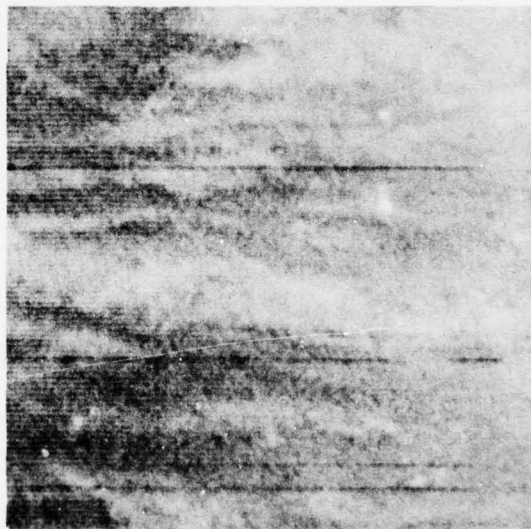


Figure 64(d). Cascaded Result
(LAGBC + MRT)
for Comparison

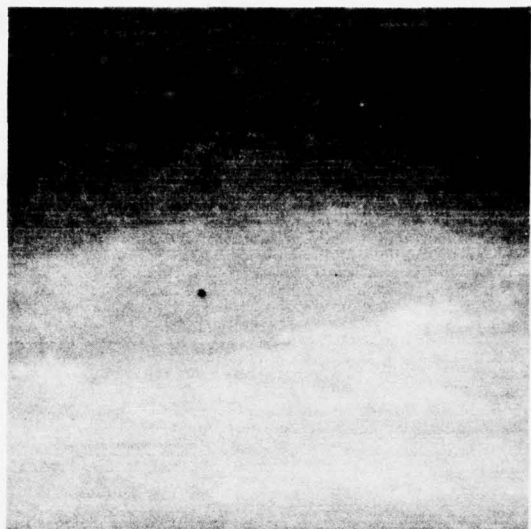


Figure 65(a). Original

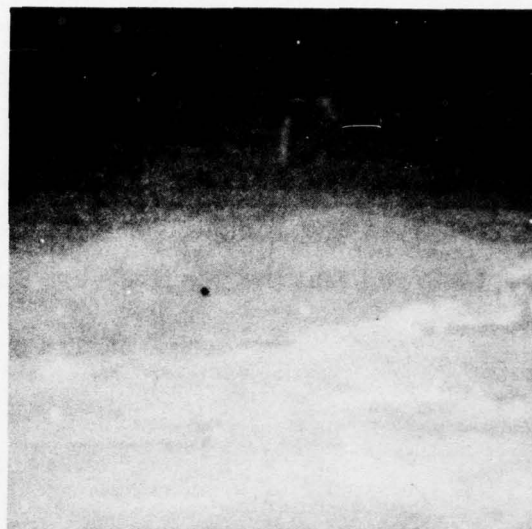


Figure 65(b). Integrated Algorithm 2

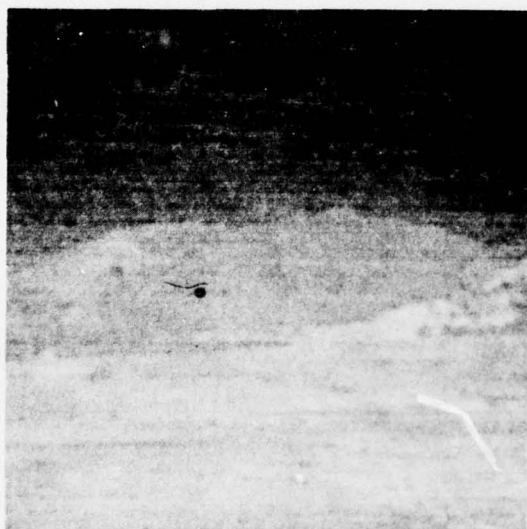


Figure 65(c). Integrated Algorithm 1



Figure 65(d). Cascaded Result
(LAGBC + MRT)
for Comparison

smoothing filter. This was because this MRT enhancement filter performed most consistently with images of varying amounts of noise. Integrated image enhancement algorithms and their results with FLIR imagery were also presented. Limited resources prevented exhaustive evaluation of these algorithms and comparison with the cascaded images. But the limited test set showed that the results were comparable.

SECTION VIII

IMAGE ENHANCED FLIR PERFORMANCE MODEL

This section develops a methodology for predicting the performance of an image enhanced FLIR. The performance of FLIRs has been traditionally characterized by the minimum resolvable temperature (MRT) curve, which is the minimum temperature difference of a standard thermal target observable when viewing the FLIR display. The MRT is measured as a function of the bar target spatial frequency (in cycles/unit angular subtense). The MRT is a function of various FLIR system components--the optics, detector, electronics, display and observer eye model. Various models have been proposed [7] and [8] to predict the MRT of a FLIR system without actually measuring it, of which possibly the most comprehensive is the NVL model developed by the Army Night Vision Laboratory. In the NVL model, the various system components are modeled by their linear system transfer functions and the MRT is expressed as a function of these component transfer functions. Computer programs have been developed by NVL and others that implement this model. Given the system specifications they output the predicted MRT. This provides a powerful design tool in FLIR system design and evaluation. The performance prediction model is part of the NVL thermal imager model; it predicts the static probability of target acquisition (detection) and recognition as functions of the MRT of the system, the target angular subtense and the target temperature contrast.

In this section we extend the NVL model to include nonlinear image enhancement subsystems. The resultant image enhanced MRT curve is then used

to predict the recognition and detection performance in precisely the same manner as the NVL model.

The following is a brief description of the procedure used to predict the MRT curves for image enhanced FLIRs. The basic definitions used for the MRT and the noise equivalent temperature difference (NETD) closely follow those of the NVL model [7]. But the approach to computing the MRT for the image enhanced system differs from that used in the NVL programs. The difference arises from the nature of the enhancement processes themselves: they can be nonlinear and more importantly, position variant--different points in the image may be treated differently by the enhancement algorithms.

For example, the LAGBC employs different local gains depending upon the local scene variance. This is nonlinear as well as position variant. The variable width adaptive smoothing filters used for interframe MRT enhancement, although linear, are by very definition position variant. These algorithms therefore defy characterization by a linear system transfer function in the frequency domain which is confined only to linear shift invariant systems. In turn this means that their effect cannot be included in the analytical NVL program as yet another transfer function in the chain between the detector and display.

This is why we turn to Monte Carlo methods and computer simulate the entire system, to include the temperature bar targets, optics and detector MTFs, additive detector noise (from a pseudo-random number generator), electronics, the enhancement process, matched filter, and the eyeball MTF. Since we have a noise corrupted bar target image on hand we can process it by the image enhancement algorithms, in spite of position variance and non-

linearity, and estimate the output signal-to-noise ratio (SNR) on the enhanced image. In this way, we avoid having to characterize the image enhancement processes by a linear shift invariant transfer function.

THE METHODOLOGY

Figure 66 shows a basic image enhancement FLIR model with the component MTFs of the optics, detector, electronics, display, and the eye (hypothesized as a matched filter [7]). The image enhancement box represents the nonlinear position (or shift) variant element in the model. In its absence we could find the system transfer functions for the signal and noise predict the MRT in a more or less closed form, as done in the NVL computer program. Here, the following Monte Carlo simulation approach takes into account the enhancement nonlinearities.

Figure 67 is a flowchart of the simulation outline. In the present simulation process, standard four-bar target images (length $L = 7W$, the width) are generated in the computer for each frequency. Spatial filters corresponding to the optics and detector MTFs are then applied to the resultant image, blurring the targets. Gaussian noise of a given variance σ^2 is then added from a pseudo-random number generator to simulate the detector noise. This is further filtered spatially to simulate the electronics; it undergoes image enhancement and is further filtered by the display MTF and the matched filter postulated in [7]. The signal and noise in the resultant image are then measured. The output signal is measured as the average peak-to-peak signal between successive peaks and valleys in the final bar pattern. The output noise is measured as the mean square difference between the noisy and noiseless processed images. The output SNR is then computed, and the

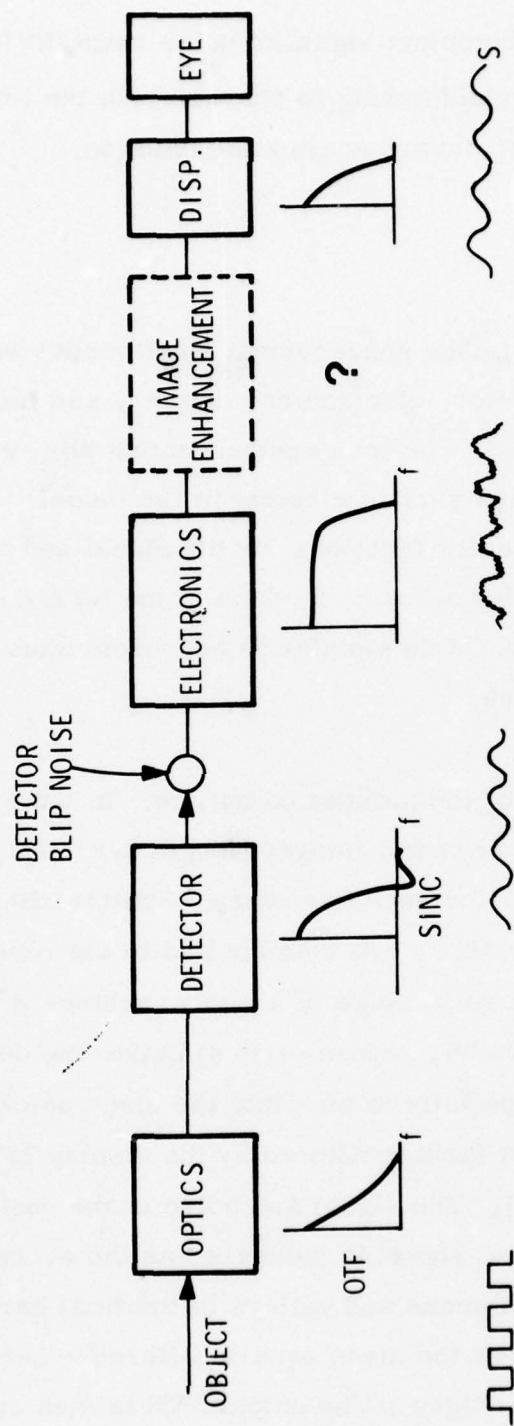
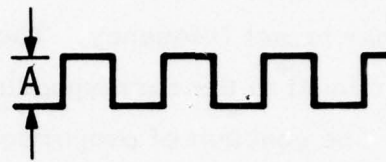


Figure 66. Image Enhanced FLIR Model

- GENERATE BAR PATTERN
 - BLUR LINE BY LINE WITH
DETECTOR AND OPTICS BLUR
 - ADD WHITE NOISE
(CONSTANT)
 - ENHANCE
 - MATCHED FILTER (EYE)
 - MEASURE SIGNAL AT OUTPUT S
 - MEASURE NOISE AT OUTPUT N
(STANDARD DEVIATION)
 - OUTPUT S/N RATIO = ?
(2.25)
- REVISE A, GO TO STEP 1



$$A' = \frac{A}{S/N} \times 2.25$$

Figure 67. Simulation Outline for Predicting the MRT

whole process is iterated by changing the input bar pattern amplitude (for a given input noise variance σ^2) until the desired output SNR is obtained for each bar target frequency. Then the MRT (ΔT) at that spatial frequency is proportional to the corresponding bar pattern amplitude used in the simulation. The constant of proportionality is a function of the NETD of the FLIR system under consideration. We now give a brief description of each step of the simulation procedure.

Generation of Bar Pattern

For the simulation, we need bar patterns of varying spatial frequencies. It is convenient to use a number of pixels to represent even the smallest bar targets ($2W = D$, the width of the detector) so that the effect of sampling can be ignored as in the NVL model (in the scan direction). All frequencies are normalized to the detector sinc function cutoff; i.e., $f_0 = \frac{1}{D}$, where the detector width D is specified in terms of number of pixels in width. The bar targets generated then have periods ($2W$) of 4, 6, 8, 10, ... pixels, corresponding to spatial frequencies of $\frac{f_0}{0.8}$, $\frac{f_0}{1.2}$, $\frac{f_0}{1.6}$, $\frac{f_0}{2}$, The length of the target is made $7W$ and a margin of 20 pixels is allowed all around. The targets have an amplitude of A with a uniform background (Figure 67).

The Optics and Detector Blurring

The generated bar targets are subjected to spatial filtering by the optics and detector point spread functions (PSF). Since only the horizontal MRT is being considered, we can perform this one-dimensional filtering on a line-by-line basis across the target and ignore the effect on the ends of the

bars (longitudinal degradation). The detector PSF is a Rect function (Figure 68) of width D pixels (5) and the optics PSF can be either a Gaussian or an "exact" PSF obtained by inverse transforming the frequency domain optical transfer function. Figure 68 also shows a Gaussian optical PSF, which in turn corresponds to a Gaussian OTF in the frequency domain.

The actual filtering of the targets can be implemented either spatially as a convolution; or in the frequency domain by the FFT, forward transforming the PSF and the target image lines, multiplying the corresponding coefficients, and inverse transforming. When care is taken to pad the functions with zeros, the circular convolution becomes equivalent to the linear convolution obtained by direct spatial domain filtering. In the present system we have elected to do this filtering by spatial convolution. The convolution y of an input image line x with a PSF h is given by

$$\text{by} \quad y(m) = \sum_{k=-M}^M h(k) x(m-k), \quad m = 1, \dots, N \quad (1)$$

The discrete $2M + 1$ point PSF $h(k)$ is determined by the shape of the OTFs and the sampling frequency chosen for the simulation. In the present example, the detector was assumed to be five pixels wide. The corresponding PSF (a real function) then would be as follows

$$h_D(k) = \frac{1}{5}, \quad k = 0, \pm 1, \pm 2$$

The optical transfer function for the diffraction limited optics is approximated by a Gaussian

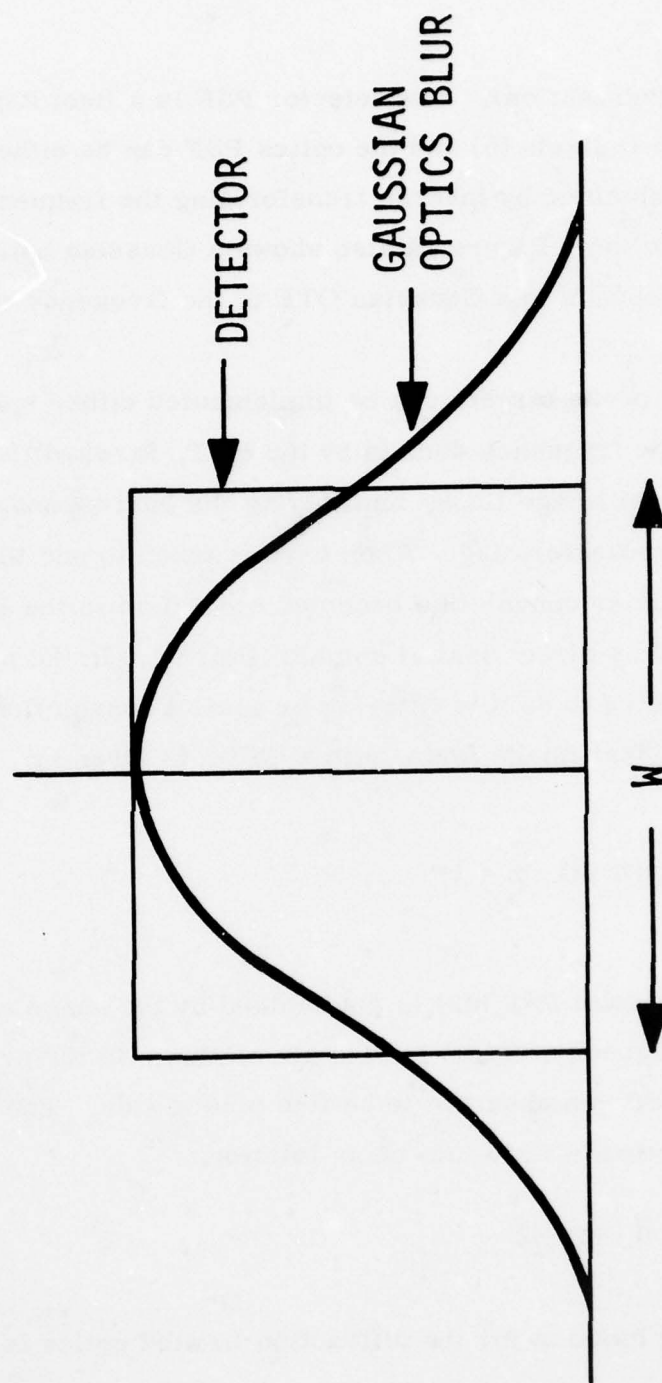


Figure 68. Detector and Optics (Gaussian) Proportions Assumed

$$h_o(m) = \alpha \text{Exp} \left[-\frac{m^2}{2\sigma^2} \right], m = 0, \pm 1, \dots \quad (2)$$

where α is determined by requiring

$$\sum_{M=-\infty}^{\infty} h_o(m) = 1 \quad (3)$$

This normalization is done merely to assure that the "gain" of this filter over a uniform field would be unity. Although the Gaussian (and the PSF of any real life optical transfer function) is infinite in extent, it decays rapidly with increasing m and is truncated at a value of m for which it is less than the quantization (8 bit) error. In the present example, the OTF was assumed to be a Gaussian with $\sigma' = \frac{1}{D} = f_o$, the detector cutoff frequency. This results in a Gaussian PSF with $\sigma = \frac{1}{\sqrt{2\pi}} D$ in the spatial domain.

The two filters, the optics Gaussian MTF and the detector Rect function, are applied in cascade using the same filter structure.

Detector Noise

The detector noise added to the degraded bar patterns is assumed to be uncorrelated in this simulation, i.e., white, possessing a flat spectrum. The additive noise is generated by a pseudo-random number generator yielding Gaussian random numbers with zero mean and unity variance. The desired noise variance σ^2 is obtained by scaling these random numbers by σ , before adding them to the targets. Note that if f_s is the sampling frequency in the simulation,

$$\sigma^2 = \int_0^{f_s} S_o(f) df = S_o(f) f_s \quad (4)$$

where $S_o(f)$ = the noise power/unit frequency interval. We will see that this relation will be useful in relating the results of the simulation to the MRT of the total system.

Electronic Transfer Function

The electronics, including the detector pre- and post-amps, can be modeled by an equivalent low pass RC circuit with a 3 dB cutoff frequency $f_{3 \text{ dB}}$, the corresponding transfer frequency is:

$$H_E^2(f) = \frac{1}{1 + (f/f_{3 \text{ dB}})^2} \quad (5)$$

This filter, which is causal, can be simulated digitally by a first order recursive low pass filter operating on the scan lines in the scan direction (for a serial scan FLIR). This filter is given by the following recursive relation:

$$y(m) = \delta x(m) + e^{-\delta} y(m-1) \quad (6)$$

where

$$\delta = 2\pi \frac{f_{3 \text{ dB}}}{f_s}$$

f_s is the sampling frequency.

In the above y is the filtered output and x is the input to the filter. In the example shown, however, the electronics filter option was not used; i. e.,

AD-A049 564

HONEYWELL INC MINNEAPOLIS MINN SYSTEMS AND RESEARCH --ETC F/G 17/5
AUTOMATED IMAGE ENHANCEMENT TECHNIQUES FOR SECOND GENERATION FL--ETC(U)
DEC 77

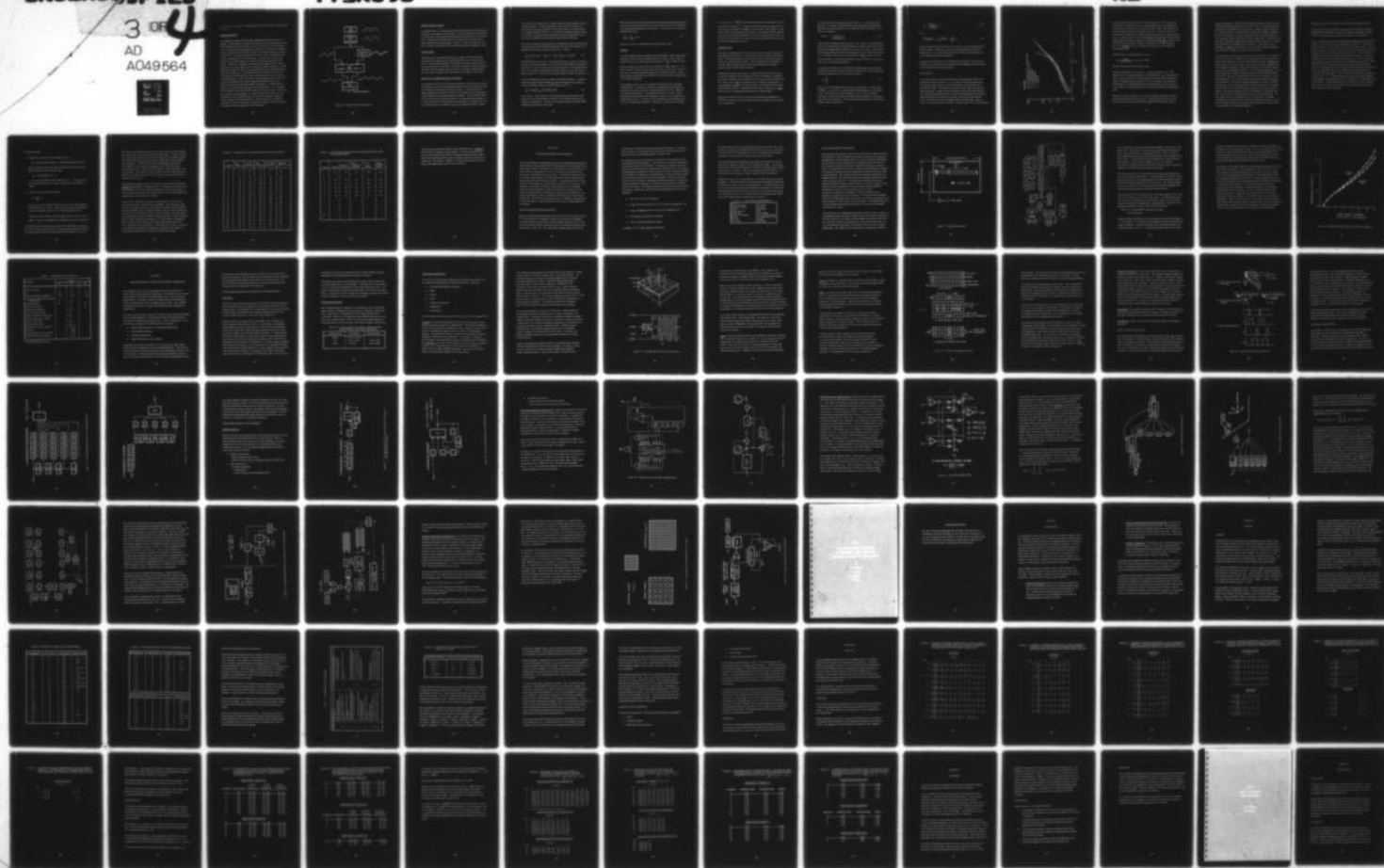
DAA053-76-C-0195

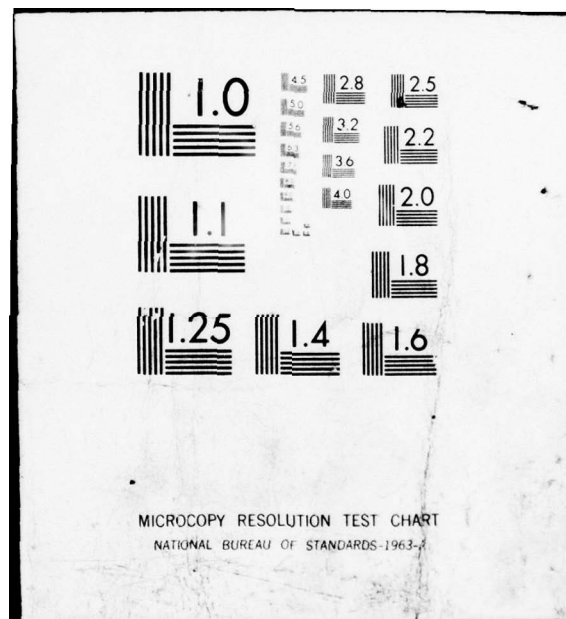
NL

UNCLASSIFIED

77SRC93

3 OF 4
AD
A049564





the electronics frequency response was assumed to be flat over the frequency.

Image Enhancement

As configured in Figure 66, the image enhancement is done just before the display. Consequently, the targets are subjected to the enhancement after having been blurred, noise corrupted and filtered by the electronics. In the simulation the image enhancement process is performed separately on the noisy blurred as well as the noise-free blurred targets as shown in Figure 69. It should be emphasized that this is the most important feature of this methodology, as we will see below. In particular, the noisy and noise-free (blurred) images are treated identically at every point in the image. Consider for example, the variable width adaptive filter used for intraframe smoothing: Here one of five different filters is applied at each point in the image, the filter choice depending on the scene contrast as measured at that point. In the present context the filter choice at point (x, y) would be determined from the noisy blurred image, as in a real life situation. This same filter is then applied to both the noisy and noise-free images that are being processed in parallel at point (x, y) . Figure 69 illustrates this schematically. Similarly, in LAGI the local gain G_{ij} is computed from the noisy version of the image. This gain is applied to both the noisy and noise-free blurred targets. Thus we can accommodate nonlinear, as well as shift variant processing in the simulation. As we shall see below, the "enhanced" noise-free image is used to measure the final signal level, while the difference between the "enhanced" version of the noisy and noise-free images is used to measure the noise at the final output state of the simulation.

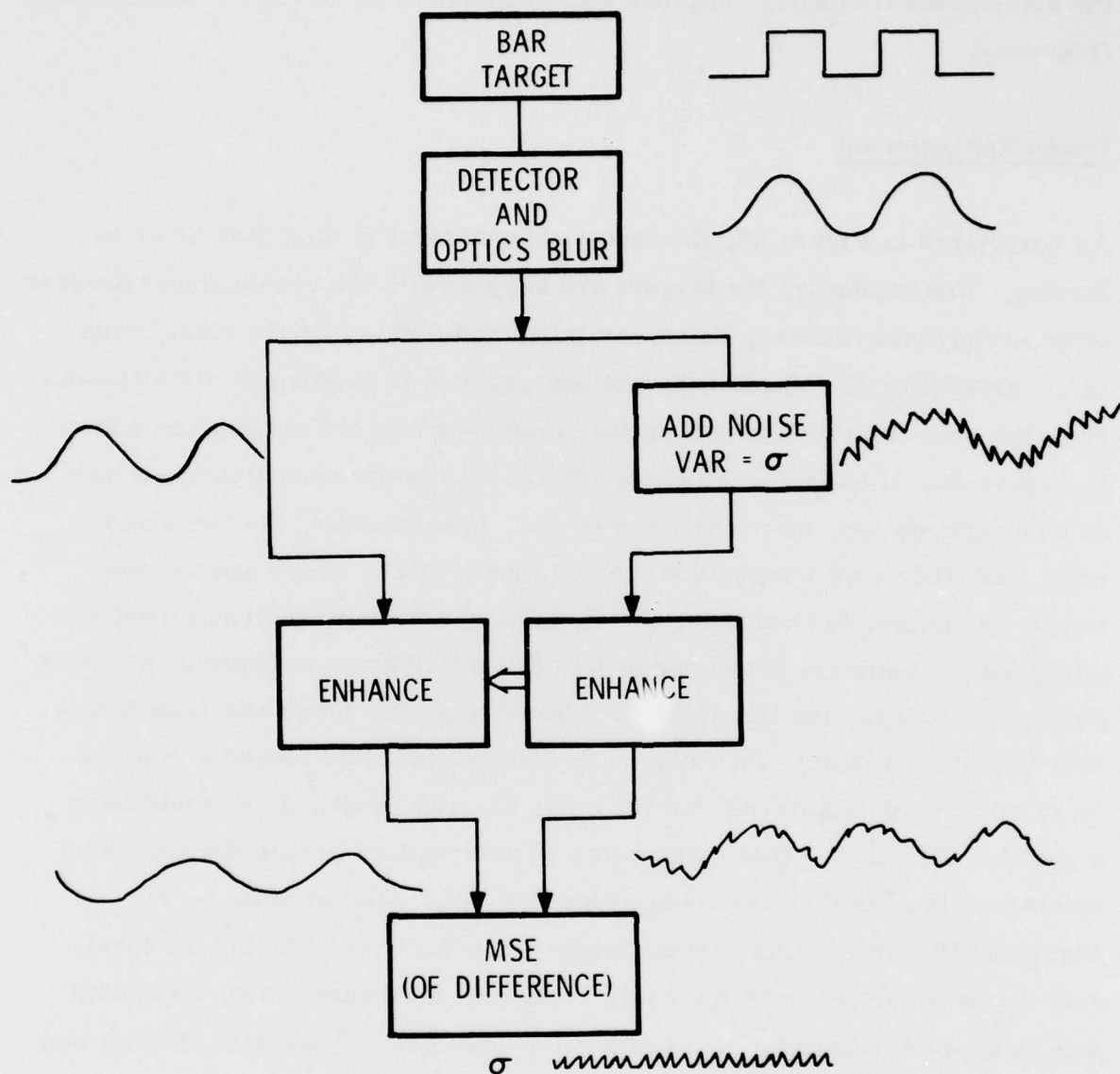


Figure 69. Measurement of Output Noise

Display Transfer Function

The display MTF is the blurring caused by the finite spot used in the reconstruction of the image on a CRT display. This is a function of the spot size and shape and is best simulated by a spatial convolution (also in one dimension). The shape most often assumed is Gaussian and this filter can be easily applied by the existing filter structure. In the examples here, however, we did not include the display transfer function for simplicity's sake.

Matched Filter

The final step in the processing of the bar targets is the matched filter (unity gain) to simulate the matched filtering done by the visual process. This is a Rect function applied (convolved) on a scan line basis to both the noisy and the noise-free version. The width of the Rect function is equal to the width of the bar targets being simulated. This again is applied using the same filter structure as the original optics and detector blur filters.

Measurement of Signal and Noise at the Output

As mentioned earlier, the output signal amplitude is measured on the noise-free blurred target which has undergone exactly the same processes except for the addition of the noise as the noisy version. Figure 69 shows the signal amplitude measured on the noise-free version as the average peak to trough difference on successive peaks and valleys. This is also averaged over all the scan lines across the target for a more consistent estimate. It is easy to see that, for the case with no enhancement, the signal at the output measured here is precisely the same as predicted by the NVL model,

even though it was measured on a target instead of being computed using the transfer functions. The only difference is that in the NVL model it is assumed that the signal amplitude is the first harmonic of the squarewave for a computational convenience. To the extent this approximation is valid, the two ways of computing the output signal should be in agreement in the case with no nonlinear or shift variant enhancements in the chain.

The noise is measured at the variance of the difference over the target area, between the final (match filtered) versions of the noisy (X_n) and noise-free processed (X_o) targets as shown in Figure 68. That is,

$$\sigma_{out}^2 = \frac{1}{N} \sum [X_n(i, j) - X_o(i, j)]^2 - \left[\frac{1}{N} \sum \{X_n(i, j) - X_o(i, j)\} \right]^2 \quad (7)$$

An equivalent but easier way of computing the above, at least in the quantized case of 8 bits for example, is to form the 512 level difference histogram of $[X_n(i, j) - X_o(i, j)]$ for all (i, j) and find the variance of the resultant histogram. In this way the expensive squaring operation need be done only 512 times instead of N^2 times (N can be as big as 500) as implied in Equation (7).

In the no enhancement case when all the processes are linear, it is easy to see that the output noise variance σ_{out}^2 given by Equation (7) is similar to the one predicted by Equation (A43) in Reference [1]; i. e.,

$$\sigma_{out}^2 = \int S(f) H_{ELECT}^2(f) H_d^2(f) H_W^2(f) df \quad (8)$$

Note, however, that because the matched filtering has been only in the x direction, the averaging effect of the noise in the vertical direction is not implicit in Equation (8). This has to be explicitly taken into account while computing the preceived output SNR.

The perceived output SNR is obtained by accounting for the spatial averaging of the noise in the vertical direction of the bars (vertical matched filtering) and the temporal averaging over six frames. This results in the following multiplicative factor.

$$\left(\frac{S}{N}\right)_p = \frac{S}{\sigma_{out}} \cdot 6L \quad (9)$$

where L is number of independent lines along the bar target.

Iteration

In a purely linear system, suppose that an input signal A with an input noise variance of σ^2 gave an output signal-to-noise ratio of K . Then, to get an output S/N of K_o , we would need an input signal $A' = \frac{K_o}{K} A$, when the noise

variance remains the same. (This is constant for a given detector.) Therefore, if the S/N measured above turned out to be K , to get a $S/N = 2.25$ at the output, we could compute the value of A' necessary as above. This A' when properly scaled gives the $\Delta T(MRT)$ for the system under consideration at that spatial frequency.

Unfortunately, this is not true of nonlinear systems such as the image enhanced FLIR. Therefore, an iterative procedure has to be followed to determine the proper A to get the requisite output $S/N = 2.25$. Figure 67 illustrated this procedure. An intelligent choice of A is made, the bar pattern generated and the entire simulation process is run to measure the output $(S/N)_p$ ratio, K . If this is 2.25, the process is terminated and A is the desired input signal amplitude. Otherwise, a prediction for the input

A' is made as $A' = A \frac{2.25}{K}$ and the simulation process is repeated with a new bar pattern of intensity A' . In general, this will not give an output (S/N) of 2.25 because the system is nonlinear. But this $(S/N)_p$ will be closer to the desired value than in the previous iteration. The process is again repeated with successively refined estimates of A' until the procedure converges. It has been our experience that the desired $(S/N)_p$ is obtained after three trials with every example we have tried.

Scaling the Axes

Note that by repeating the above process for every bar target frequency we have obtained an MRT curve. But the axes of this curve are not in familiar units. The horizontal axis is scaled in terms of W , the number of pixels across one cycle of the bar targets. The vertical axis is the amplitude of the input target pattern relative to the standard deviation of the input detector noise σ , added in the simulation.

Scaling the horizontal axis is easy. Remember that the OTF of the optics used in the simulation was expressed in terms of the detector dimensions. In addition, the detector was assumed to be a certain number of pixels wide (D). Hence, a bar target of full cycle width of W pixels corresponds to a spatial frequency of $\frac{D}{W} f_o$ mrad/cycle where f_o is the detector sinc function cutoff frequency (which is usually the system cutoff). Note that $f_o = \frac{1}{\text{IFOV}}$, where IFOV is the instantaneous field of view in milli-radians.

Scaling the vertical axis is a bit trickier and requires knowledge of the NETD of the detector and the particular definition employed in the measurement of the NETD.

Let us suppose that the detector noise is white, i. e., it has a flat power spectrum. If a target of sufficiently large dimensions and a temperature difference ΔT gave an output of A at the detector output, the NETD of the system is defined by

$$\text{NETD} = \frac{\Delta T}{A / \int S(f) H_E^2(f) df} \quad (10)$$

where $S(f)$ is the detector noise spectrum. Now, $H_E(f)$ is the equivalent low pass RC circuit frequency response designed to represent the actual FLIR and post-amp high frequency behavior, as in the NVL model. Some authors take a slightly different view [8] that $H_E(f)$ used for the NETD measurement is the "standard" filter with a cutoff frequency of $\frac{1}{2 T_d}$ where T_d is the detector dwell time. With either definition, $H_E(f)$ is treated in exactly the same way in the following discussion.

In the above simulation let us assume that the white noise variance added was σ_o^2 , held constant over all the trials for different bar target frequencies. Then the noise power per unit bandwidth is given by

$$S_o = \frac{\sigma_o^2}{f_s} \quad (11)$$

where f_s is the sampling frequency used in the simulation. In the current example, $f_s = 5f_o$ where f_o is the detector cutoff frequency ($f_o = \frac{1}{D}$ cycles/mrad). For a given bar target frequency f , let A be the input target amplitude (in the same units as σ_o) required in the simulation. We need to relate the A to the actual ΔT . We do this using the definition of NETD of the system, shown in Equation 10.

$$\Delta T = \frac{\text{NETD}}{S(f)H_E^2(f)df} \cdot A \quad (12)$$

where

$$\int_0^\infty S(f)H_E^2(f)df = S_0 \int_0^\infty H_E^2(f)df = \frac{\sigma_0^2}{f_s} f_R \quad (13)$$

The above assumes white noise $S(f) = S_0 = \sigma_0^2$ and f_R is the cutoff frequency of the "standard" reference R_C low pass filter $H_E(f)$ used in the NETD measurement (or, equivalently, the cutoff of the RC circuit modeling the pre- and post-amp low pass behavior).

Equations (12) and (13) give the scaling factor to scale the vertical axis of the MRT curve obtained by the simulation and the frequency axis is normalized with respect to the detector cutoff.

AN EXAMPLE

Let us assume that the FLIR in question has a detector/optics blur in the proportions shown in Figure 67. (It is trivial to change this proportion). The additive noise variance σ_Q in the simulation is, for example, 66.67 (in arbitrary units). Given this we can arrive at the MRT curve using the above simulation. Figure 70 shows a family of such curves obtained with the various enhancement schemes in the chain between detector and display. The horizontal axis is scaled in terms of f_0 , the detector sinc function cutoff. Borrowing from the second generation FLIR example, Table 2 in Section 9, the detector IFOV = 50 μ rad. Hence $f_0 = \frac{1}{50} \times 10^6 = 20$ cycles/mrad.

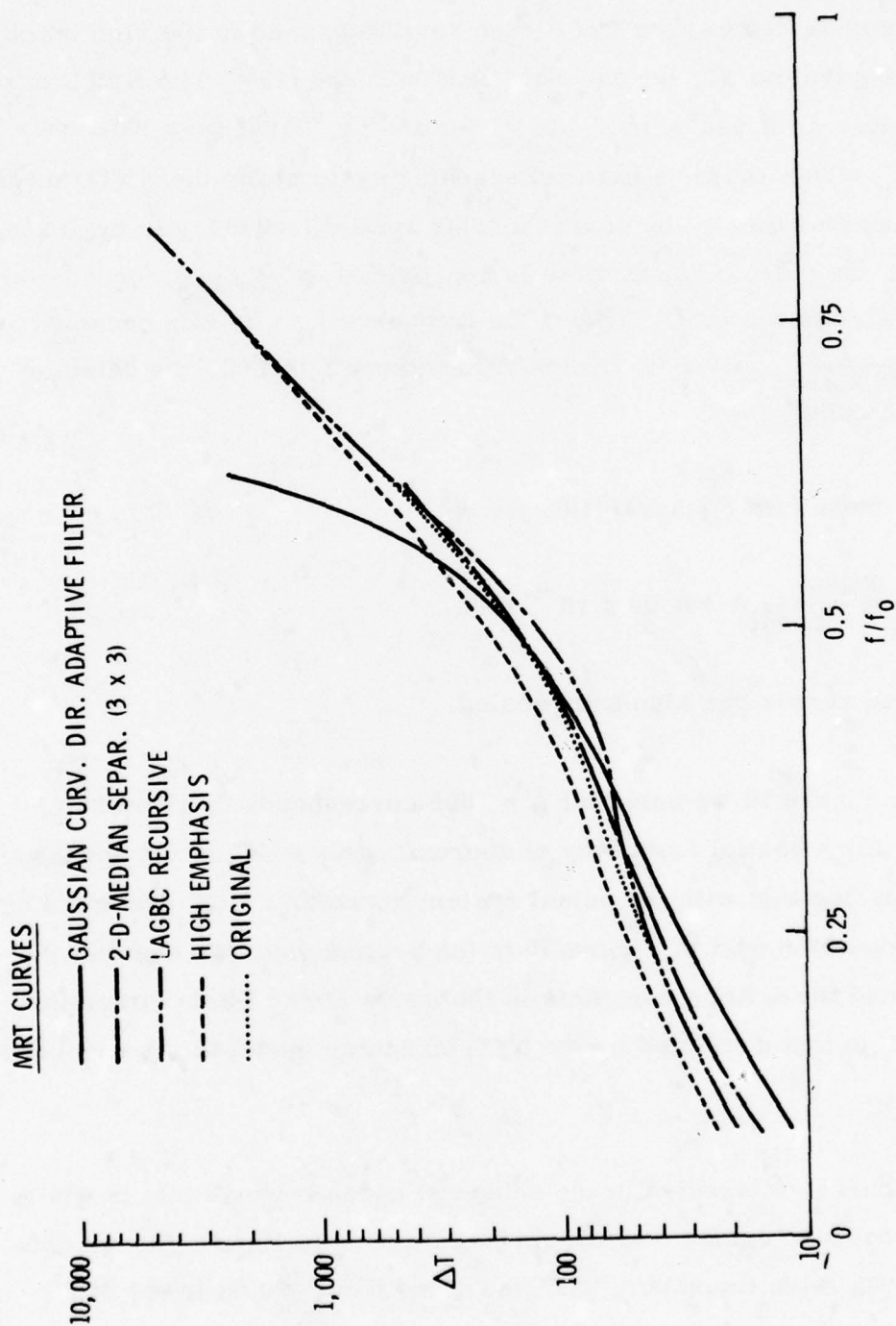


Figure 70. Predicted MRT Curves Using the Monte Carlo Approach
 f_0 = detector cutoff

So we have scaled the horizontal axis in absolute terms. The vertical axis in Figure 68 is scaled in terms of A , the amplitude used in the simulation. To find the equivalent ΔT , we use Equations (12) and (13). The NETD of the system is given as 0.034°K in Table 2, Section IX. Assume a reference bandwidth $f_R = 30 \text{ kHz}$ (for a parallel scanning system) for the NETD measurement. To convert this to the corresponding spatial frequency in cycles/mrad, assume that one detector dwell time is approximately $\frac{1}{60} \times \frac{1}{800}$ seconds for 800 active IFOVs per scan line. Hence, the frequency $f_R = 30 \text{ kHz}$ becomes $30 \times 10^3 \times \frac{1}{60 \times 800} f_o = 0.63 f_o$ cycles/mrad where $f_o = 1/D$, the detector sinc function cutoff.

Substituting these into Equation (12), we get

$$\Delta T = \frac{0.034}{66.67 \times .63} A = 8.09 \times 10^{-4} A^\circ\text{K}.$$

Now the vertical axis has also been scaled.

Referring to Figure 70 we note that $A = 1000$ corresponds to $\Delta T \approx .8^\circ\text{K}$. This occurs for a spatial frequency of approximately $0.7 f_o$ in the unenhanced case. Comparing this with the actual system horizontal MRT predicted by the NVL computer model in Figure 75 in the next section, we see that the MRT shape and the scaling constants in the above Monte Carlo simulation is very close to that predicted by the NVL computer model in the unenhanced case.

We are, of course, interested in the enhanced cases because that is where the above procedure comes into its own. Referring to Figure 70, we note that the nonrecursive Gaussian MRT smoothing filter yields lower MRT

at higher frequencies. High frequency emphasis appears to perform poorly at all frequencies except close to the system cutoff. It is conceivable that it might have crossed the unenhanced curve at a higher frequency (closer to f_0). But the simulation was not carried so close to the singular frequency. The two dimensional 3×3 median filter also improves the low frequency MRT apparently without degrading the high frequency MRT. But this was because the smallest bar target was four pixels in width, and the 3×3 median filter therefore did not degrade this bar target. In the absence of oversampling, we can expect the median filter to degrade the higher frequency MRT somewhat, although not as much as linear filters.

The surprising result, however, comes from the LAGBC curve in Figure 70 where we see that it is asymptotic to the nonenhanced curve at both the low and the high frequencies. But in the mid-frequency range, it improves the system MRT, by about 25 percent. We attribute this to the adaptive nature of the LAGBC formulation. Recall that the local gain is inversely proportional to the local standard deviation. At low bar target frequencies, the bar target amplitude is small in comparison with the local standard deviation due to noise. Hence the LAGBC gain is set to unity (do nothing). At very high frequencies, also, the local standard deviation is higher, although it is contributed to now by the bar targets themselves. Here again LAGBC does nothing. In the mid-frequencies, the signal amplitude is comparable to noise, and the size of the local area is big enough to enhance the bar targets. The local gain gets a big boost, amplifying the bar target amplitude. Note that this boosts the noise also at bar target frequencies. But the measured noise is integrated over the entire spectrum, although modified by the matching filter, whereas the signal is confined to a narrow band of frequencies about the target bar pattern frequency. Hence we see an improvement in the measured SNR here.

METHODOLOGY FOR PREDICTION OF RECOGNITION PROBABILITY

Once the MRT curve is specified, we apply the NVL model to predict the probability of recognition of a military target given its angular subtense and the sensor temperature contrast with respect to its background.

Often, we would like to predict the probability of recognition from thermal images, after the image has been acquired and digitized. To do this, we need the MRT curve of the sensor from which the image is taken. We also need the absolute temperature difference of the targets and their dimensions in angular subtense in order to use the NVL model. Since the absolute temperature is often not available in the digitized image due to hardware calibration errors and the declassification process, we need to relate the intensity difference ΔI to the actual noise variance in the image σ_o^2 . Then $\Delta I / \sigma_o^2$ becomes a normalized variable with which we can go to an MRT curve. We developed [2] a way of measuring the image noise variance σ_o^2 from the image texture feature that yields a better estimate than the variance of the image itself. This was derived from the texture measure (contrast, for $\Delta = 1$). The following process can then be used to predict the probability of recognition with the NVL tables in [7] and the MRT curve for the sensor. Using the different MRT curves for the various enhancements, we can predict the probabilities of recognition if the images had been enhanced by these algorithms.

THE PROCEDURE

1. Measure the target ΔI from the original image

$$\Delta I = |\text{average target intensity} - \text{average background intensity}|$$

2. Get an estimate of the image noise using the first order difference texture statistics on the original image

$$\sigma_N = 1/2 \text{ CONTRAST } (\Delta = 1)$$

i.e., use the CONTRAST texture feature for $\Delta = 1$. This σ_N may be the value measured for that image or averaged over the class of images.

3. Compute the equivalent ΔT as follows:

$$\Delta T = \frac{\Delta I}{\sigma_N} \cdot \sigma_0$$

where σ_0 is the noise standard deviation used in the image enhanced MRT model, and $\sigma = 66.67$. Given this ΔT , find the corresponding normalized frequency (f/f_0) on the corresponding MRT curve.

Multiply the target height in pixels by $(\frac{f}{f_0})$ to get "cycles" across the target. Then get the probability of recognition from the NVL table [7].

We applied the above procedure to estimate the probability of recognition for a set of 20 thermoscope images for which we had previously measured the average target contrast and the texture feature (contrast, for $\Delta = 1$).

We assumed that these images came from a sensor with the same detector optics blur circle ratio as the example simulated above. Table 3 presents the results of this prediction for these 20 images for the original, LAGBC, high emphasis filtering, two-dimensional median filter, and the adaptive smoothing filter. Note that the predicted probabilities (interpolated) from Table 3 for the various enhancements do not differ much from the unenhanced case. However, LAGBC appears to improve the low probability images. The majority of these images were very high contrast and therefore very high on the MRT curve, where the various MRT curves are asymptotic to each other.

Comparing their results with that in Table 4, which are the corresponding measured probabilities of recognition from the evaluation phase (see Part II of this report), we note that these predicted values are not consistent with the measured values across the various images. Across enhancements the changes appear to be consistent, however.

In summary, a Monte Carlo approach to predicting the MRT of a FLIR system with nonlinear components was developed. This approach, based on the NVL model, was shown to give identical results to the conventional approach using the linear transfer functions when all the system components in the FLIR can be modeled by linear transfer functions. The advantage of the new approach is that it can handle nonlinear and space variant processes such as image enhancement processes in the FLIR system. Examples of MRT curves of a FLIR system with various image enhancements were shown to illustrate this methodology. Using these MRT curves, we attempted to predict the probability of recognition in digitized FLIR images, by estimating the signal (target contrast) to noise in these images.

TABLE 3. PROBABILITY OF RECOGNITION PREDICTED BY MODEL

Image	Non-Enhanced	LAGBC Recursive	High Emphasis	2-D Median Separable	Gaussian MRT Curv Dir
1	.97	.97	.96	.97	.95
2	1	1	1	1	1
3	.91	.91	.89	.91	.86
4	.72	.7	.70	.74	.70
5	.98	.98	.98	.98	.97
6	.88	.89	.86	.89	.85
7	.32	.34	.29	.22	.20
8	.81	.82	.78	.82	.78
9	.70	.70	.65	.70	.67
10	.87	.88	.84	.88	.84
11	.85	.87	.82	.86	.85
12	.67	.70	.63	.69	.65
13	.95	.95	.94	.95	.92
14	.41	.41	.39	.41	.37
15	.77	.79	.70	.77	.77
16	.29	.37	.25	.29	.32
17	.52	.56	.48	.54	.52
18	.22	.27	.15	.22	.22
19	.44	.50	.39	.46	.46
20	.13	.15	.11	.13	.13

TABLE 4. MEASURED PROBABILITIES OF RECOGNITION FROM THE EVALUATION PHASE

Image	Original	LAGBC Recursive	High Frequency Emphasis	2-D 5 x 5 Median	MRT Adaptive Filter
1	.39	.38	.34	.23	.24
2	-	-	-	-	-
3	.18	.27	.08	.46	.20
4	.54	.61	.57	.80	.73
5	.54	.60	.52	.50	.50
6	.93	.50	.58	.74	.36
7	.36	.56	.64	.31	.31
8					
.					
.	-	-	-	-	-
.					
20					

These results unfortunately showed little correlation with the measured probability of recognition on these images. This could be explained in part by the fact that the true MRT of the system producing these images, which have undergone several stages of preprocessing from the sensor to digital tape, was unknown. However, the Monte Carlo approach for producing the image enhanced MRT has been verified.

SECTION IX

SECOND GENERATION FLIR EXAMPLE

This section defines a representative FLIR as an example demonstrating the methodology for image enhanced FLIR design using the Image Enhanced FLIR Performance Model. Since there is a wide spectrum of DOD missions from which to draw a representative FLIR, it is necessary to establish some criterion for selecting a particular mission. The criterion chosen was that the mission should require a FLIR that allowed the fullest exercise of the Image Enhanced FLIR Performance Model. A consequence of this criterion is that the FLIR might not be optimized for a generic set of missions. For example, to exercise the resolution enhancement portion of the Image Enhanced FLIR Performance Model it is necessary to configure a FLIR where the detector/IFOV is significantly smaller than the optical spot size. Similar observations could be made about several other key FLIR system parameters. Within the constraint stated above, we define a representative mission, develop the mission parameters and present a second generation FLIR design.

MISSION PARAMETERS SPECIFICATION

Both the Advanced Attack Helicopter (AAH) and the F-16 will require FLIRs to be fully effective in fulfilling the mission for which they are designed. The requirements and constraints on FLIR designs imposed by the F-16 are greater than those imposed by the AAH due to the aerodynamic performance requirements of the F-16. For this reason a typical mission for an F-16

was selected to develop the mission parameter specification. It should be noted, however, that the differences between the two applications are not large and the same methodology applies to both aircraft.

The following scenario was selected as a basis for the development of the second generation FLIR example. The mission is to neutralize tanks and/or Armored Personnel Carriers (APCs). The scenario consists of an initial phase during which the aircraft approaches the indicated area at low altitude; a detection phase during which the pilot pops up to about 1,500 feet and scans the designated area looking for potential targets; a recognition phase during which the pilot examines several potential targets to determine the nature and priority of the targets; a deployment phase during which he releases his armament against the target(s); a damage assessment phase; and finally a return to base phase. The following salient requirements were derived from this scenario, taking into consideration typical defense levels and weather conditions.

1. Wide field of view for navigation
2. Target detection starting at 9 to 12 km from the designated area
3. Target recognition by 6 to 9 km from the designated area
4. Ten seconds to accomplish recognition
5. Clear to moderately adverse weather

In addition, a 1° ΔT target signature is assumed.

The aircraft imposes size and weight constraints on the FLIR. The size of the dome containing the gimballed FLIR optical system must be maximized to reduce drag to an acceptable level. Here it is assumed that six inches represents an upper limit for the size of the entrance aperture.

The aircraft scenario also imposes constraints on the amount of time the pilot has for scanning a monitor while searching for targets. Accordingly, the infrared scan must be displayed in a fashion that maximizes the transfer of information to the pilot. Full image enhancement will be required, as well as some sort of automatic target recognition and cueing system.

Finally, there are a set of requirements that are determined by peripheral conditions. A standard 60 field, 30 frame-per-second display compatibility is assumed for interface with existing equipment. As mentioned above, the 8 to 12 micrometer wavelength spectral band is assumed for the FLIR since the diffraction effects due to the six inch aperture will exercise the resolution enhancement portion of the Image Enhanced FLIR Performance Model. The system requirements are summarized in Table 5.

TABLE 5. FLIR SYSTEM REQUIREMENTS

Detection Range	9 to 12 km
Recognition Range	6 to 9 km
Target	Tank/APC
Target Contrast	1°C
Background	27°C ambient
Spectral Wavelength	10 to 12 micrometers
Weather	50 percent RM at 27°C
Aperture	6 inches
Frame Rate	30/second

SECOND GENERATION FLIR DESIGN

The next generation of FLIRs will have to provide both increased sensitivity and resolution in order to achieve the necessary detection and recognition probabilities at the long ranges required to minimize aircraft vulnerability. Aperture sizes will be restricted due to packaging constraints; therefore the increased sensitivity can only be achieved by increasing the detector count. The high detector count in turn requires the development of integrated focal planes where much of the signal processing and formatting is done on the focal plane. One focal plane concept particularly well suited to TV display systems is the vertically scanned, horizontally multiplexed linear array as shown in Figure 71. The focal plane consists of an array of detectors coupled to a CCD structure in such a way that the output from the detectors can be integrated for a line dwell time, parallel shifted to a multiplexing structure, and then serially read out a video stream that is directly TV compatible. Greater sensitivity can be achieved by adding more rows of detectors. The outputs of the additional detectors are added coherently to the initial row of detectors through a time delay integration (TDI) structure on the focal plane. The detailed structure of the focal plane and limits on the number of TDI stages are discussed elsewhere in this report.

A block diagram for a typical high performance FLIR is shown in Figure 72. The telescope focuses the scene radiation at the field stop that contains two reference surfaces at the upper and lower edges of the field of view (FOV). The reference surfaces are maintained at a temperature differential, ΔT , and provide known signal inputs to the detector array for gain and offset equalization. The radiation from the field stop is collimated, scanned in

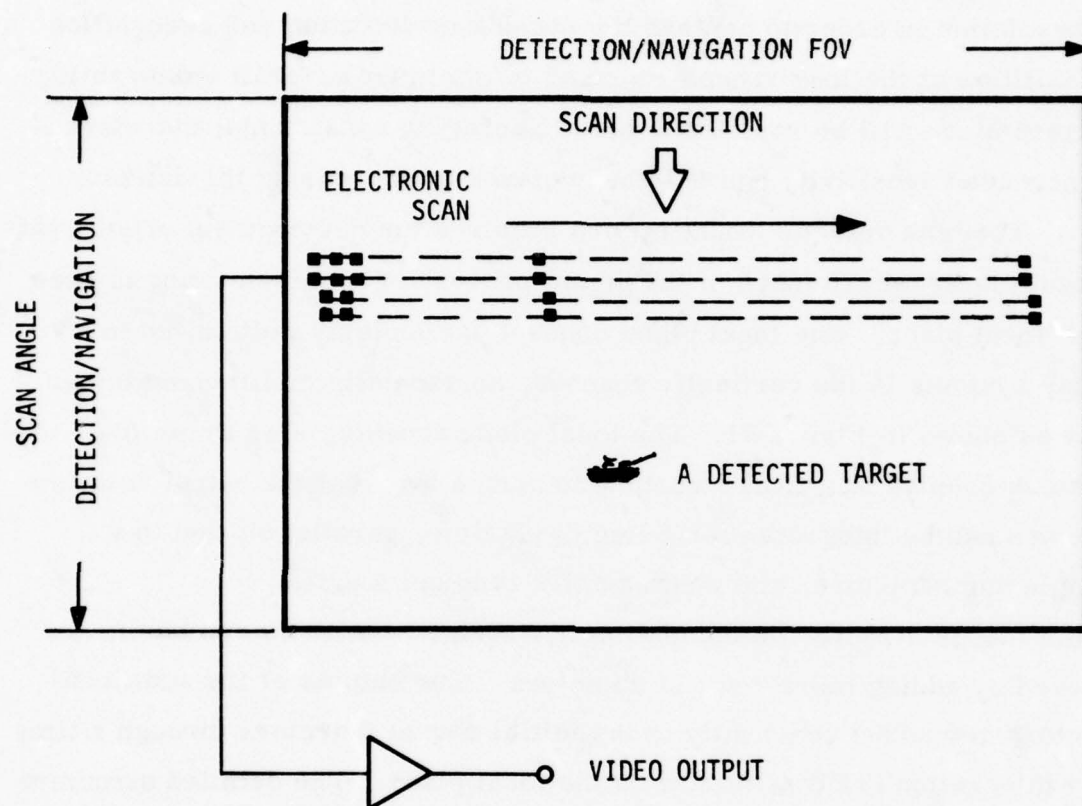


Figure 71. Focal Plane Concept

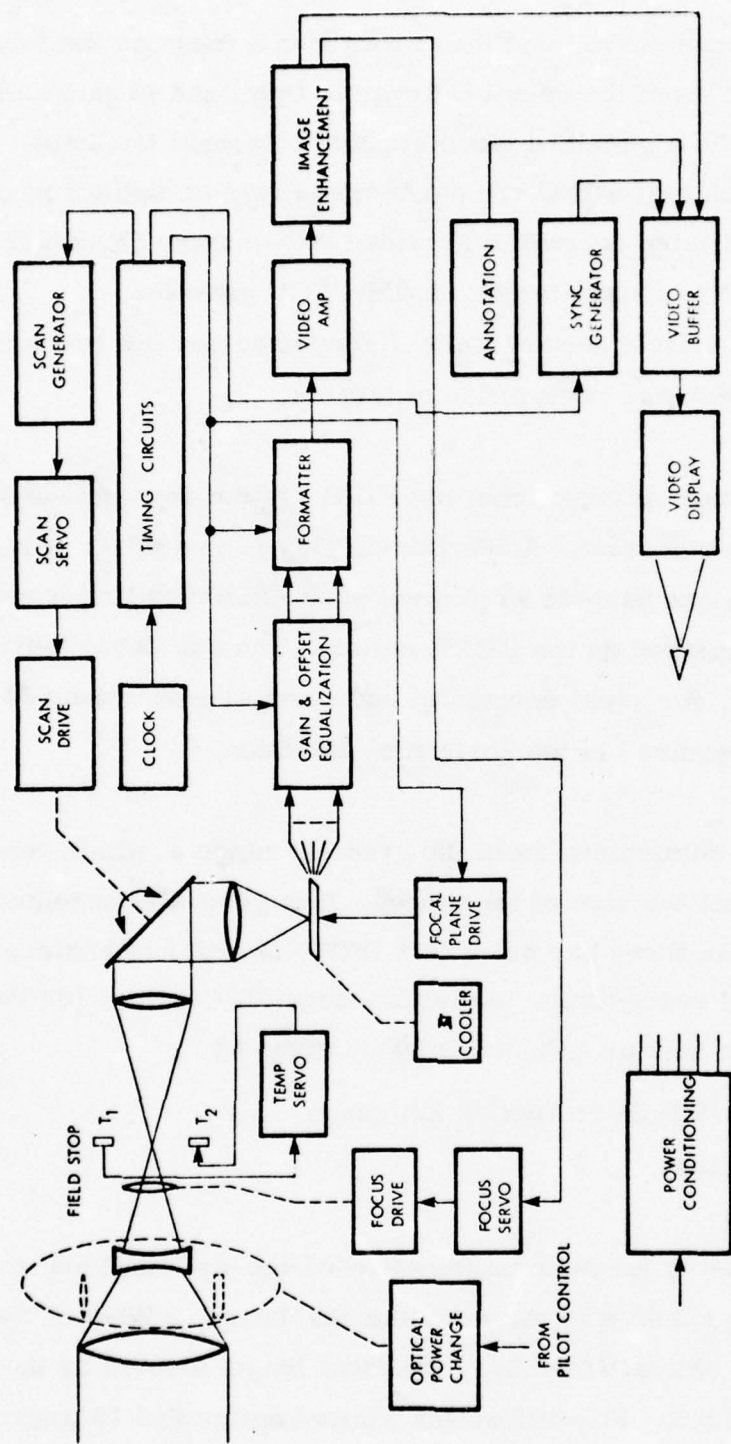


Figure 72. High Performance FLIR Block Diagram

the vertical plane by a scan servo, and then brought to a focus on the focal plane array. The output from the detector array is equalized in gain and offset, then formatted into a standard video stream. Various forms of image enhancement, such as LAGBC are performed prior to display on the CRT monitor. Overall timing control is provided by a master clock and logic circuits. Two FOVs are available. A wide FOV provides coverage for navigation while a narrow field of view provides the resolution necessary for the detection and recognition of targets.

The system parameters are derived from the FLIR system requirements and certain empirical information. A 525 line display is assumed. A high performance scan servo can achieve 90 percent scan efficiency that leads to 473 lines of scene displayed on the CRT monitor. The standard display format is 4:3; therefore, for equal horizontal and vertical resolution 630 detector elements are required in the horizontal direction.

The size of the IFOV is determined from the greatest range at which recognition is desired and from the size of the target. It is generally accepted that a target must subtend three bar pairs (six IFOV) in order to achieve 50 percent probability of recognition. Assuming about 2.7 meters for the minimum dimension of a tank or APC the IFOV is given by

$$\begin{aligned}\text{IFOV} &= (2.7 \text{ meters} / 6 \text{ IFOV} / \text{target}) / 9 \text{ km range} \\ &= 50 \text{ microradians}\end{aligned}$$

Minimum detector geometry for near future state-of-the-art focal planes is 1 x 1 mil detectors. To achieve 50 microradian resolution, a 20-inch focal length is required. The optical f/number, the focal length divided by the aperture diameter, is f/3.3. For diffraction limited optics and 10 micro-

radiation, 84 percent of the energy is contained in a spot of diameter $2.24 \lambda / \text{number}$ or 3.2 mils (just over 3 IFOV). For an "as-fabricated" wave front error of 0.1 wave, the diameter of the spot encircling 84 percent of the energy increases to about 5.8 mils. Note that the relationship of detector size to optical spot size is favorable for exercising the MRT enhancement portion of the Image Enhanced FLIR Performance Model.

The number of TDI stages required for the focal plane is determined by the desired FLIR sensitivity. For this example a MRT of 0.1 degree at the IFOV frequency in the horizontal direction is required to detect a target T of 1° with an atmospheric transmission of 0.1 (50 percent RH at 27°C). The predicted performance of the FLIR is computed using the NVL Static Performance Model for thermal viewing systems as modified by Honeywell to model hybrid focal planes. For this FLIR example, an effective transmission of 55 percent and a photovoltaic detector area resistance product (RoA) of 40 are assumed. An RoA of 40 assumes near photon limited performance. Given the above assumptions, one TDI stage is required to achieve an MRT of 0.1 at the IFOV frequency of 10 cycles per milliradian. Figure 73 gives the MRT as a function of spatial frequency. The complete performance prediction program output is given in Appendix A. Table 6 lists the nominal FLIR system parameters along with acceptable limits on some of the parameters for parametric trade-off analysis.

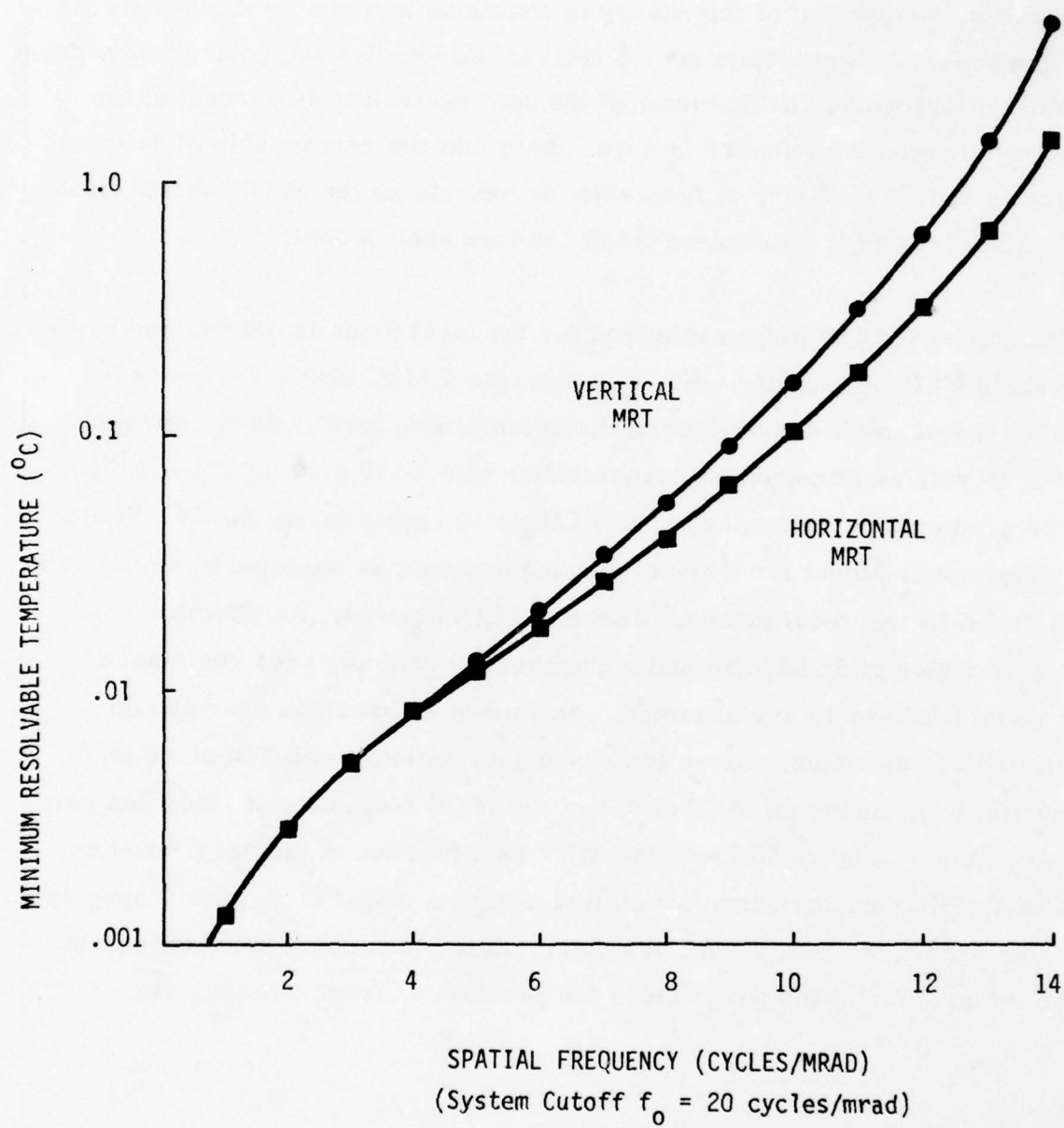


Figure 73. Predicted FLIR System MRT Using NVL Programs

TABLE 6. FLIR SYSTEM PARAMETERS

Parameter	Value		
	Min	Norm	Max
Aperture (inches)	5	6	7
f/#	4.0	3.3	2.9
Spectral Wavelength (microns)	8-11	9-12	8-14
Optical spot size for 84% encircled energy (mils)	3.2	5.8	
Transmission (%)	40	55	70
IFOV (microradians)		50	75
Field of view (degrees)		1.8 x 1.4	
Frame rate (per second)		30	
Number of detectors in parallel		630	
Number of detectors in series		2	
Detector size (mils)		1 x 1	
Detector RoA ($\Omega\text{-cm}^2$)	4	40	
Detector D^* (f/3.3) ($\text{cm Hz}^{\frac{1}{2}}\text{W}^{-1}$)		2.8×10^{11}	
NETD ($^{\circ}\text{K}$)		0.034	
Horizontal MRT (10c/mrad) ($^{\circ}\text{K}$)		0.10	
Vertical MRT (10c/mrad) ($^{\circ}\text{K}$)		0.16	

*3 bar recognition from 6 km

SECTION X

IMPLEMENTATION OF IMAGE ENHANCEMENT ALGORITHMS

Under Phase I of this contract Honeywell analyzed the performance of several FLIR image enhancement algorithms. We selected the most promising as subjects of the Phase II implementation study. We not only used performance as a selection criterion, but also considered the possibility of future implementation before making the final choices. This provided reasonable certainty that a design could be effected before beginning the design phase.

This section describes the results of the Phase II study and provides designs for the selected algorithms. Honeywell's Systems and Research Center performed the study during the period May 1, 1977, to September 15, 1977. The section is divided into the following major subsections:

- Systems Requirements and Technology Considerations
- Focal Plane Partitioning
- Hardware Subelements
- Image Enhancement Circuit Designs

Processing of FLIR data in real time at rates suitable for video display imposes several requirements on the processing hardware. The first subsection summarizes these and also surveys the technologies suitable for use in the hardware. The second subsection examines the optimum location--either on or off the focal plane--for each of the image enhancement functions.

Certain types of filtering operations are required in several of the more promising algorithms. We can utilize this commonality to develop versatile hardware subelements; these are discussed in the next subsection. The fourth subsection briefly summarizes the most promising algorithms and presents the circuit designs for these algorithms.

SYSTEM REQUIREMENTS AND TECHNOLOGY CONSIDERATIONS

Introduction

Real time or near-real time processing of FLIR data at video rates requires that the hardware be able to handle both video sampling rates and a wide dynamic range. The sampling rates requirements are between 5 and 30 MHz, the exact value is a function of the horizontal resolution desired and the number of lines (either 525 or 875) in the output display unit.

Dynamic range of the impact data can be as high as 90 to 100 dB, but a typical display provides only 20 to 30 dB. Level-shifting, thresholding, and antiblooming functions can provide up to a 20 to 25 dB compression of input dynamic range; then LAGBC and high emphasis filtering can reduce the dynamic range to that which the display can handle. For data processing beyond the focal plane we will assume a dynamic range requirement of 60 to 75 dB for the contrast enhancement function and 45 to 60 dB for the resolution restoration and MRT enhancement. The latter range is somewhat larger than that of a display, but we must perform averaging and interpolation functions which require expansion to a larger dynamic range and subsequent compression for display. For example, interframe smoothing of eight frames require an additional 18 dB of dynamic range in the processor,

although we can restrict the output data to the original dynamic range by normalizing the result (divided by the number of frames).

Nearly all of the image processing algorithm studies require at least one two-dimensional linear filtering operation. Operations of this kind require access to previously sampled points in the serial video data stream. That access requires delay/storage of individual pixels, entire lines, and entire fields. We will first examine the delays and storages required, then perform a trade-off analysis of various hardware alternatives.

Video Delay and Storage

The horizontal resolution required in a FLIR system with TV-compatible output is 500 to 1,200 pixels; the exact choice is a function of the desired MTF. This number, in conjunction with the video sampling rate of 5 to 30 MHz, allows us to calculate the range of delays required to provide pixel, line, and field delay. Table 7 summarizes the pixel storage and time delay required to provide delays of one pixel, one line, and one field.

TABLE 7. VIDEO STORAGE AND DELAY REQUIREMENTS
FOR ONE PIXEL, ONE LINE, ONE FIELD

Delay Function	Time Delay	Storage (Pixels)
Pixel	0.03 to 0.2 μ sec	1
Line	30 to 63 μ sec	500 to 1200
Field	30 msec	128K to 500K

Technology Considerations

There are seven basic types of calculations required to implement any of the image processing functions Honeywell has studied. These are:

- Filtering (fixed-weight or adaptive)
- Logic
- Memory
- Delay
- Arithmetic Operations
- Multiplexing
- Amplification

We will discuss each briefly and indicate the merits of possible approaches.

Filtering--The conventional approach to high speed filtering (other than simple band pass functions) is the digital filter. Considerable flexibility is gained both by the ability to synthesize arbitrary response characteristics and by the possibility of digital control of the filter to allow adaptive filtering. Unfortunately, however, the number of operations per second required at video rates can be quite high. For example, a 16-point, one-dimensional transversal filter requires 16 additions and 16 multiplications per pixel. For even a 525 line system, the processing rate required is about 1.3×10^8 operations per second--which is very difficult to achieve with any widely available technology. GaAs logic may, however, evolve into useable state within the next several years.

Two analog approaches can meet the speed requirements, however: charge transfer devices (CTDs) and surface acoustic wave (SAW) devices. The two largest classes of CTDs are the charge-coupled device (CCD) and the bucket-brigade device (BBD). CCDs offer higher performance (generally higher charge transfer efficiency), but BBDs have the advantage of processing simplicity and hence, cost. CCDs, however, offer an important advantage--they can be readily utilized as transversal filters. The operation of a CCD split-electrode filter (the most common type) is shown in Figure 74. The dynamic range of such filters is 65 to 70 dB and the maximum clock frequency is about 20 MHz with conventional (i.e., non-peristaltic) technology. A limitation is dark current; operation above $+30^{\circ}$ C is usually not practical without some form of cooling.

The second analog filter concept is the SAW device. These structures use propagation of the input signal along the surface of the device as an acoustic wave. Filtering is performed by a series of interleaved metallic "fingers" of various lengths; this leads to interference--either constructive or destructive--for the various frequency components in the incident wave. The output wave form is thus a filtered waveform. SAWs can easily handle the data rates video processing requires but the achievable dynamic ranges are somewhat limited (40 to 50 dB). In addition, the millisecond delays required for line and field processing are too long for SAW devices.

Adaptive filtering can be done quite easily in the digital world, although the speed limitations mentioned earlier still apply. Of course, the data must be converted from analog to digital in either fixed-weight or adaptive filtering schemes. Since the tap weights are binary words,

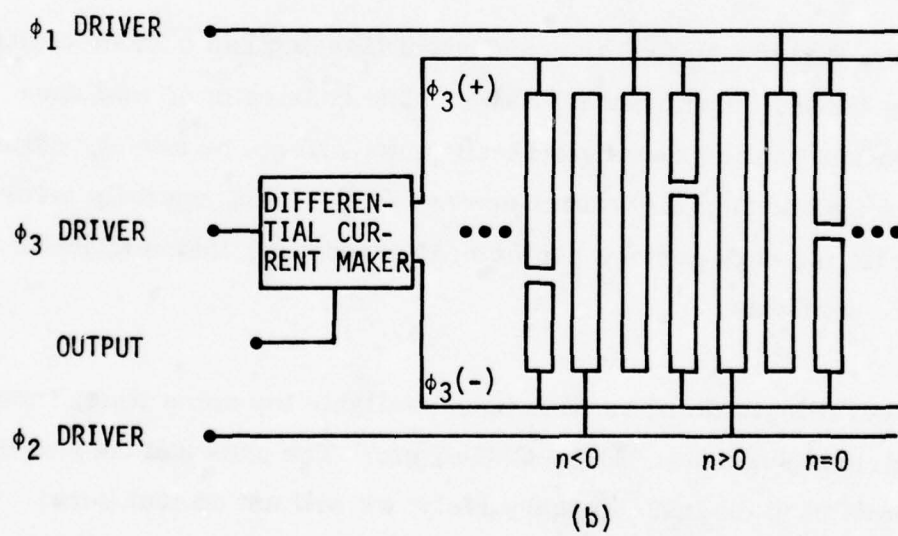
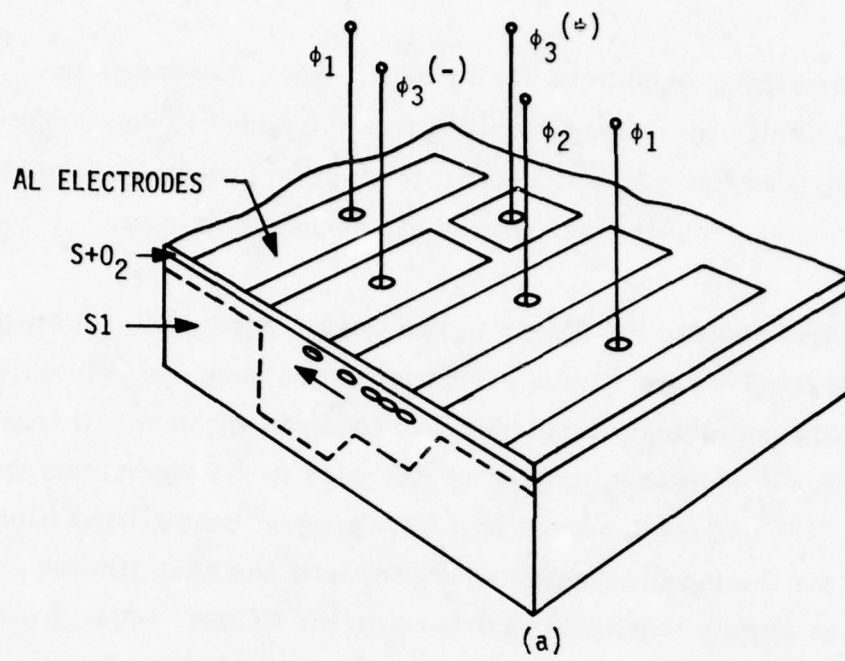


Figure 74. CCD Split-Electrode Transversal Filter

changing any or all tap weights is not difficult. Now, however, the number of arithmetic operations required must be added to the number of additional operations for adaptive control to give the processor loading--an even more severe requirement than in the nonadaptive case.

CCDs can be used as adaptive filters in two ways. First, we can switch among several fixed-weight filters. This technique requires only current technology, although clearly the method is not useful if there is a choice of many filters. Fortunately, this is not the case in the algorithms we will discuss. The second approach is a fully programmable CCD filter, although thus far the dynamic range of such filters has been limited to 40 dB or less because of gate threshold variations in the CCDs. Several methods for varying the tap weights of a surface acoustic wave (SAW) filter are currently available, or switching among fixed-weight structures is also possible.

In summary, either adaptive or nonadaptive filtering can be done digitally, with analog CCDs, or with analog SAWs. The constraint of real time video operation implies that digital techniques will not be useful, while dynamic range and flexibility work against SAWs. We generally favor CCDs despite the dark current problem; if necessary, thermoelectric coolers are available.

Logic--Several logic families have been available for some time, including TTL (in various versions), ECL, CMOS, etc. The pros and cons of these families have been discussed exhaustively; we will not do that here. Suffice it to say that TTL offers more standardized parts at lower cost than the others, and that Schottky versions of TTL are compatible with video processing rates. CMOS and some of the newer logic technologies

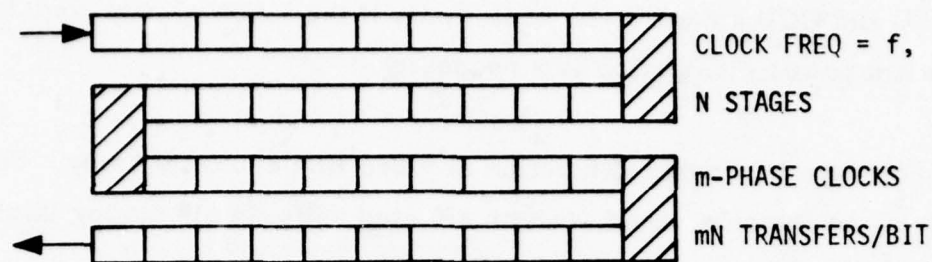
such as I^2L and CCD appear to be useful only if the logic is very complex and minimum power dissipation is a necessity.

Memory--Storage is a critical function in video line and field delay. Single pixel delays, on the other hand, can be effected with simple analog latches or CTDs, so we will not discuss them further.

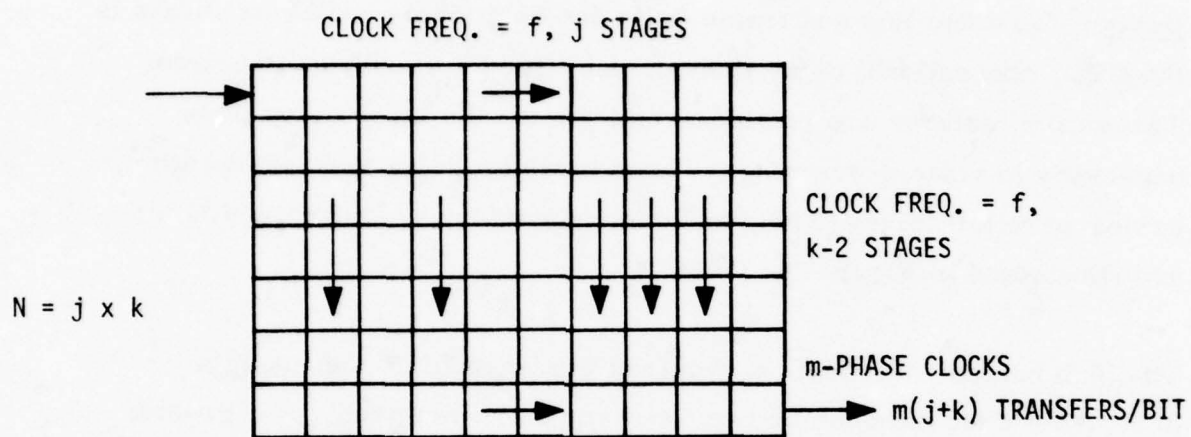
Delay--For video line and frame delay the only viable analog candidate is the CTD, conventional delay lines do not offer the flexibility required. Three architectures are in widespread use for the large delay lines necessary in video processing. These three organizations--serpentine, series-parallel-series (SPS), and demultiplex-linear-multiplex (DLM)--are illustrated in Figure 75.

The high number of transfers required in a serpentine approach is acceptable when only one or two line delays are required. For greater delays, however, the charge transfer inefficiency of the CTD can severely degrade image MTF and dynamic range. The serpentine structure is easiest to fabricate, however; corner turning is a requirement because of the near-rectangular chip size.

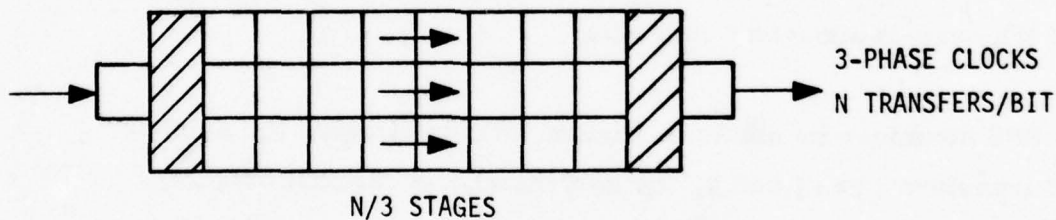
A SPS structure as shown in Figure 75(b) minimizes the number of transfers when j and k , the horizontal and vertical extents, are chosen nearly equally. The only problems with this structure are that 1) the parallel clocks run at a frequency of f/j , where f is the input clocking frequency; and 2) the transition from series to parallel clocking can result in periodic noise spikes at the parallel clocking frequency. The clocking for an SPS structure, therefore, is



(a) SERPENTINE



(b) SERIES-PARALLEL-SERIES (SPS)



(c) DEMULTIPLEX-LINEAR-MULTIPLEX (DLM)

Figure 75. N-Stage Video Delay with CTDs

quite complex. Not only must extra clocks (versus a serpentine structure) be provided, but the timing must be precisely controlled. The SPS structure is still the structure of choice in very long delay applications.

The final structure is illustrated in Figure 75(c). The so-called DLM structure transfers charge packets in adjacent registers on different clock cycles. This scheme provides a total number of transfers that is between the serpentine and SPS numbers. Only one set of clocks are needed for the CCD structures, but complex input and output structures are necessary.

In summary, the DLM structure is most advantageous for intermediate (i.e., five or six) numbers of line delays. For very large numbers of line delays, the loss due to charge transfer inefficiency is critical; we should use the SPS architecture. For small numbers of line delays the simplicity of the serpentine structure is best.

The number of pixels required for field delay is too large for current analog delay line technology. Field storage can be done digitally, however, by performing A/D conversion on the input analog serial data stream, demultiplexing, and performing digital-to-analog conversion on the multiplexed output data. The converter and multiplexer technologies are now available to do processing at video rates. In addition, CCD 64K serial memories that operate at 5MHz are available; only two to eight memories are required. Only 16 to 64 total memory chips are required, therefore, for storage of an entire video field.

Arithmetic Operations--The merits of CCD filters for performing sum of product operations were noted earlier. Both analog and digital modules are available to perform other operations (i.e., addition, subtraction, multiplication, division) at video rates. Special functions such as absolute value are also possible using CCD technology. Figure 76 indicates one way in which the absolute value of the difference of two signals can be extracted with CCDs. Assume that signals V_A and V_B are present on floating gates G1 and G2, respectively. When MOSFET C1 ties G1 to G4, charges ($\propto V_B - V_A$) will be injected into the potential well under G4 if $V_B > V_A$. If $V_B < V_A$, then connecting G3 to G4 with MOSFET C2 will also inject the difference signal $V_A - V_B$ into the well under G4. We have now calculated the absolute difference. Using this circuit, analog operation is possible at video rates.

Multiplexing--The small size and low power dissipation of CCDs make them ideal for the multiplexing of focal plane signals. In addition, the CCDs provide higher speed than more conventional alternatives such as CMOS.

Amplification--Again, numerous MOS and bipolar designs are available in chip form.

FOCAL PLANE PARTITIONING

The image enhancement algorithms described earlier in this report fall into four broad classes: DC level shifting, contrast enhancement, MRT enhancement, and resolution restoration. Such functions as antiblooming control and DC level shifting can be readily incorporated into the detector/multiplexer unit cell. DC restoration, however, requires line buffers for each line on the focal plane; it is best performed on a serial data stream off the focal plane.

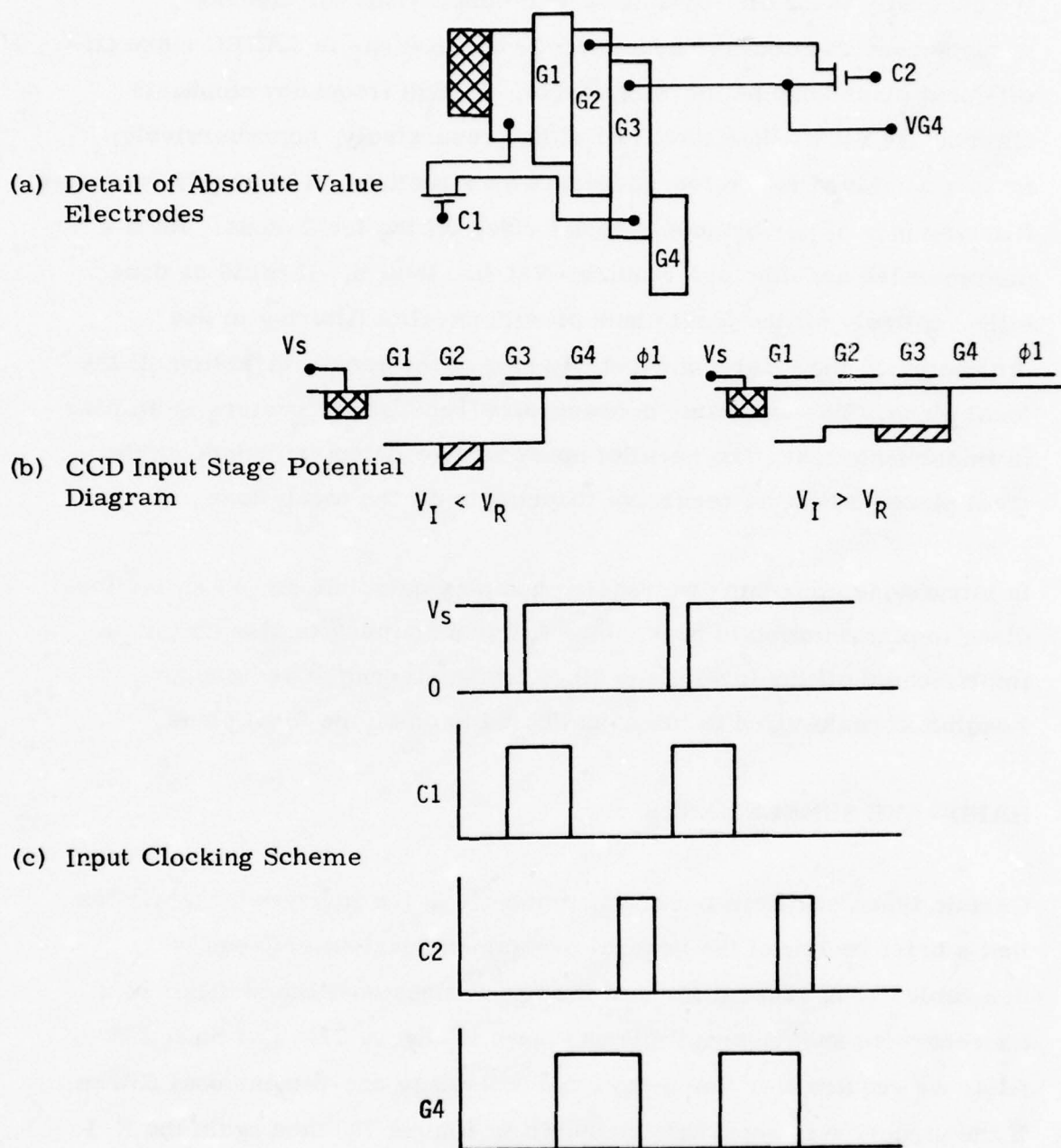


Figure 76. Absolute Value Scheme with CCDs

We generally favor off-focal plane implementations for contrast enhancement functions. The numerous calculations in LAGBC make an off-focal plane scheme more tractable. In high frequency emphasis filtering we can realize the filter either recursively, nonrecursively, or in a combined recursive/nonrecursive structure. The purely recursive filtering is best performed on serial video off the focal plane. An $N \times N$ nonrecursive architecture requires $N-1$ line delays. It could be done either entirely off the focal plane or with parallel filtering in one dimension on the focal plane and filtering of the second direction off the focal plane. The separable nonrecursive/recursive structure is similar in implementation. The parallel nonrecursive filtering is done on the focal plane, while the recursive filtering is off the focal plane.

In intraframe smoothing we require complex calculations, so an off-focal plane implementation is best. Interframe smoothing is also better implemented off the focal plane since frame storage is required. Resolution restoration is too complicated to do on the focal plane.

HARDWARE SUBELEMENTS

Certain types of filters occur so frequently in the Honeywell algorithms that a brief review of the general two-dimensional filter types is desirable. The general form of the two-dimensional linear filter is nonrecursive and nonseparable as shown in Figure 77. For an $N \times N$ filter we require $N-1$ line delays and N N -stage one-dimensional filters. If the structure is separable as shown in Figure 78, then again the $N-1$ line delays are required, but now we need only two N -stage linear filters.

2-D NON-RECURSIVE (NON-SEPARABLE)

$$y(i, j) = \sum_i \sum_j a_{ij} x(i, j)$$

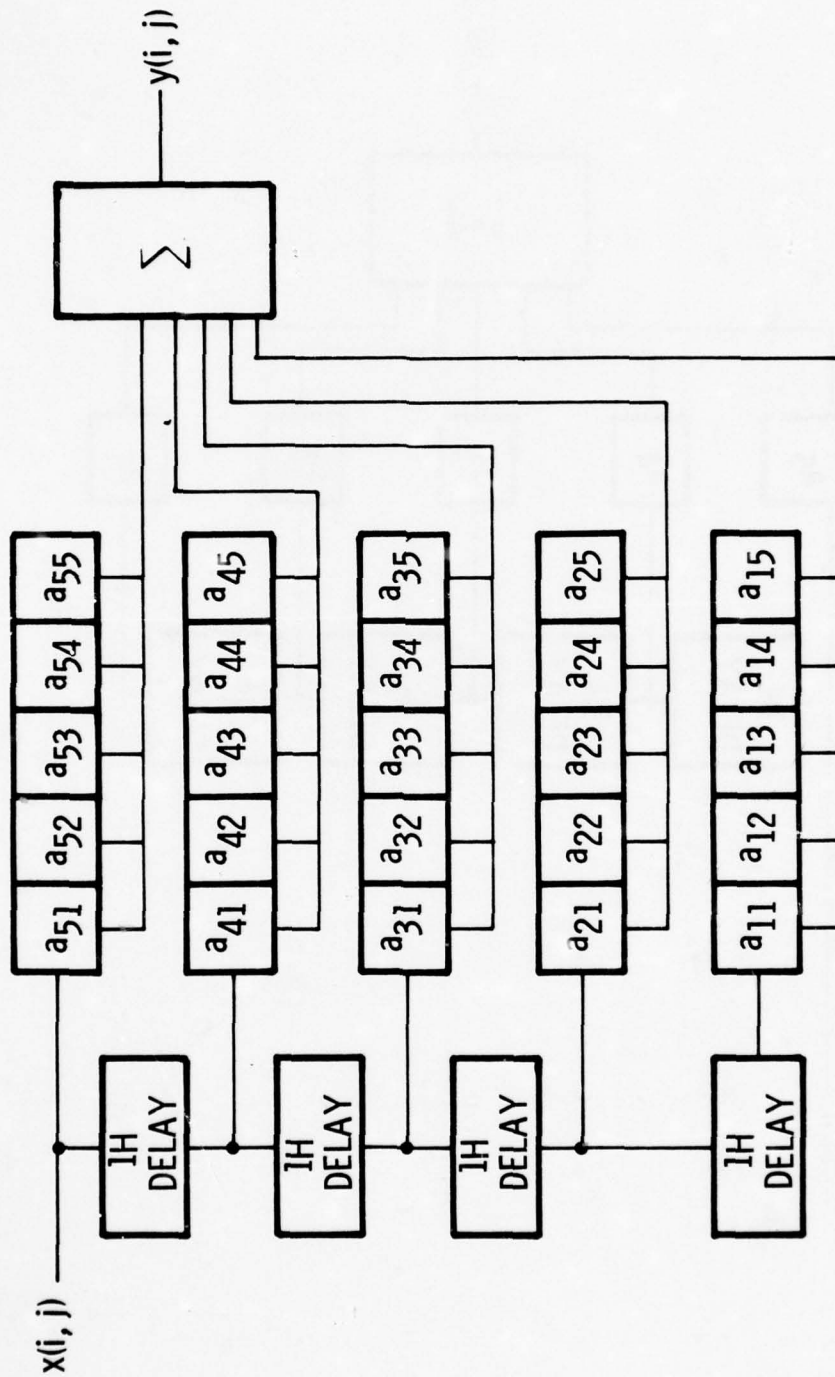


Figure 77. Two-dimensional Nonrecursive, Nonseparable Filter

2-D NON-RECURSIVE (SEPARABLE)

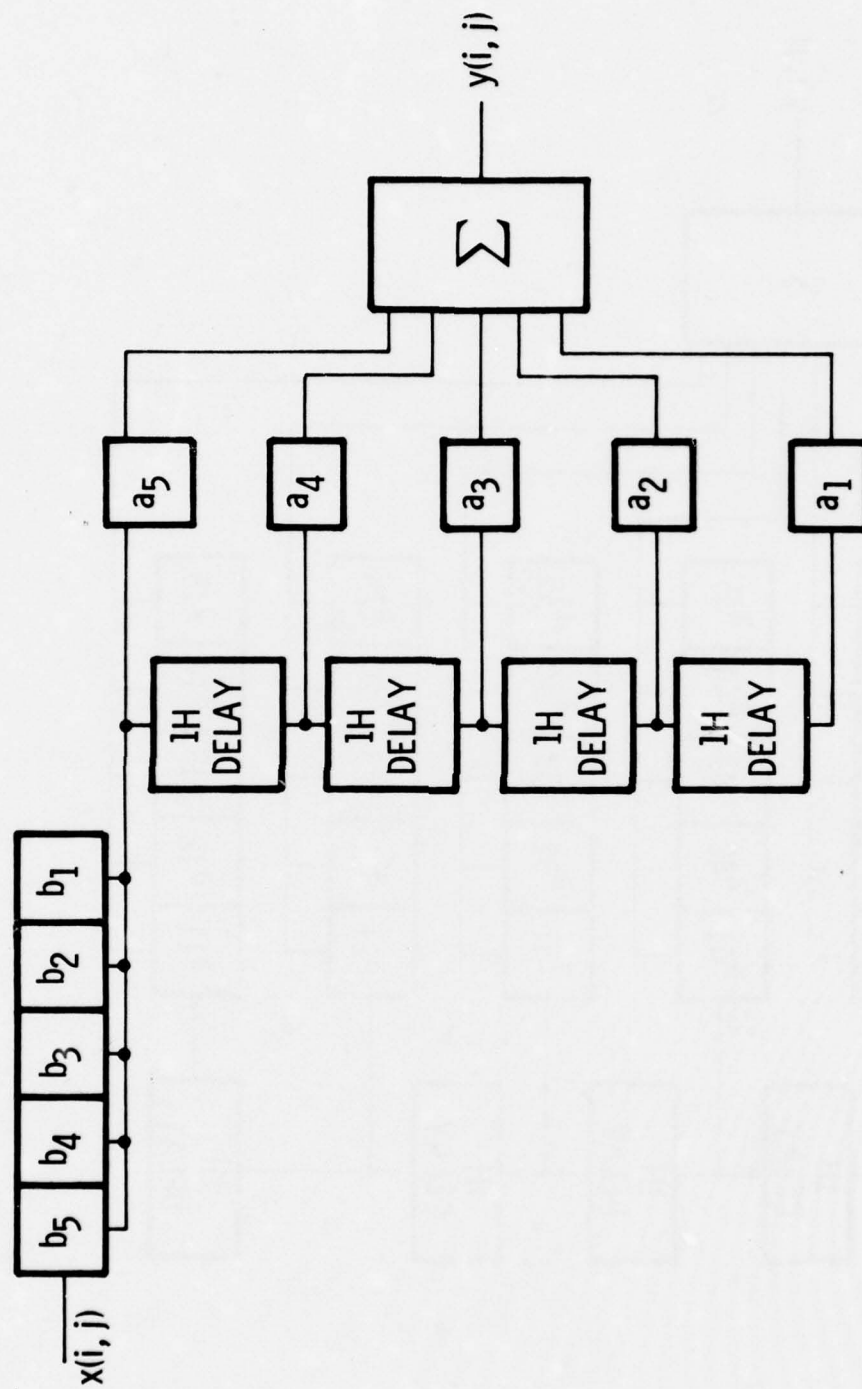


Figure 78. Two-dimensional Nonrecursive, Separable Filter

The really significant savings in hardware complexity, however, are when recursive filtering is added. A two-dimensional filter with one-dimensional recursive and the second nonrecursive is shown in Figure 79. The non-recursive filter is again a linear N-stage filter, while the one-dimensional recursive structure has a single line delay and two tap weights. The fully recursive two-dimensional filter again requires a line delay as shown in Figure 80, but only four tap weights are necessary.

IMAGE ENHANCEMENT CIRCUIT DESIGNS

Algorithm Selection

Before we present the implementations of the various algorithms, we shall briefly review the algorithms favored for implementation. Since DC level shifting function can be easily done on the focal plane, our primary emphasis here is on contrast enhancement, MRT enhancement, and resolution restoration. We will implement algorithms in each of these three areas as presented below:

- Contrast Enhancement
 - High frequency emphasis filter
 - Recursive local area gain/brightness control (LAGBC) filter
- MRT Enhancement
 - Intraframe smoothing
 - Median filter
 - Local curvature-directed adaptive filter

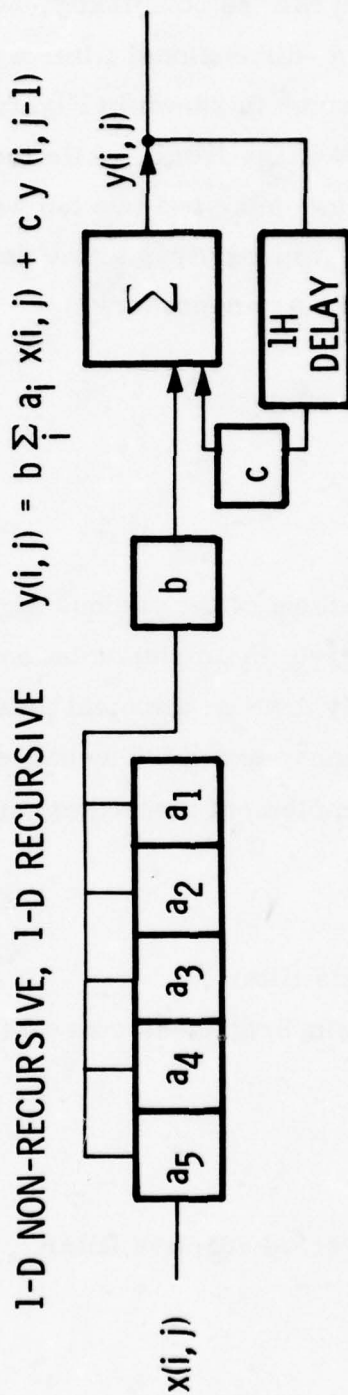


Figure 79. Two-Dimensional Filter, One-Dimensional Nonrecursive,
One-Dimensional Recursive

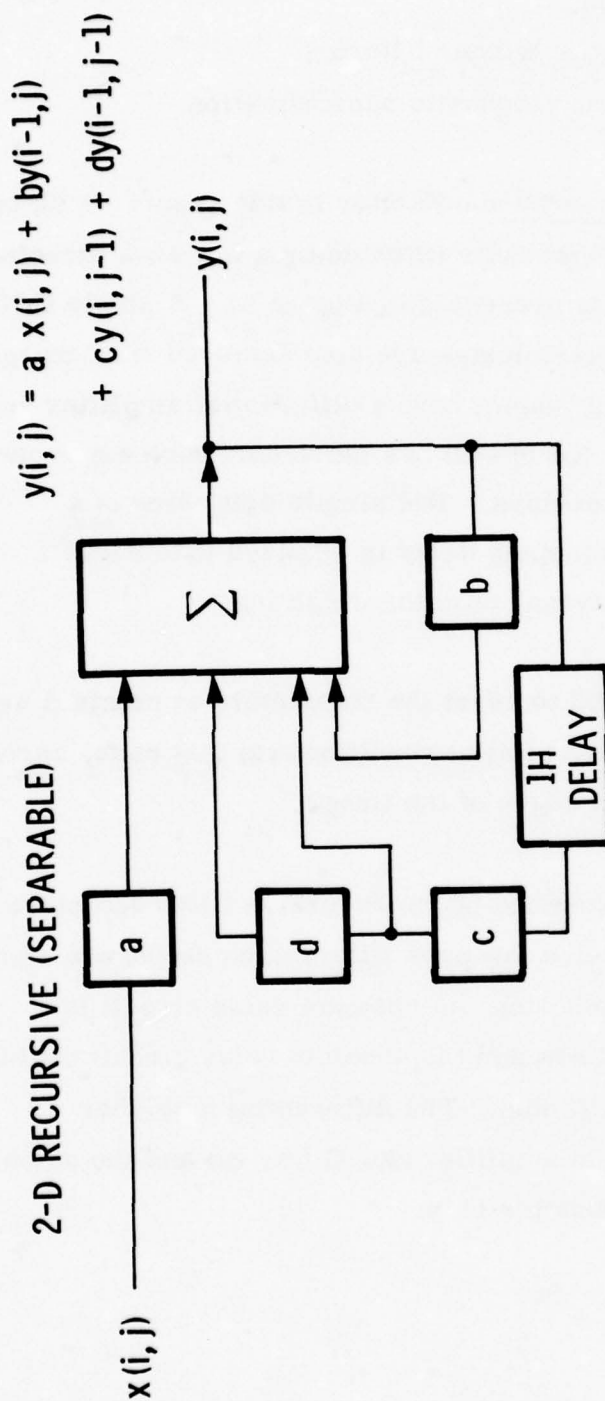


Figure 80. Two-dimensional Recursive, Separable Filter

- Resolution Restoration
 - Focus restoration with Wiener filtering
 - Superresolution using stochastic approximation

Contrast Enhancement Implementation--Earlier in this report we discussed the merits of realizing a high-emphasis filter using a low pass structure. A design for such a structure is presented in Figure 81. A single CCD line delay is required, while pixel delays are also achieved with charge-transfer devices. In the circuit shown here a differential amplifier sums the tap weights. Incidentally, Honeywell has fabricated such a structure using commercially available devices. The simple delay line is a 910-stage CCD unit, while the tapping delay is realized with bucket brigade devices which have external resistor weighting.

The counter network is provided to reset the transistors at points A as shown. This is necessary so that initial conditions (in this case, zeroes) are provided at the left and top edges of the image.

In Figure 82 we present the extension of the recursive filter structure to the LAGBC filter. Two recursive low pass filters such as the one shown in Figure 8 are required. In addition, an absolute value circuit is necessary. Both recursive filters and the absolute value circuit could be incorporated on a single CCD chip. The differential amplifier (for $I_{ij} - m_{ij}$), the variable-gain amplifier (for $G(m, \sigma)$) and the output summing amplifier would be done off-chip.

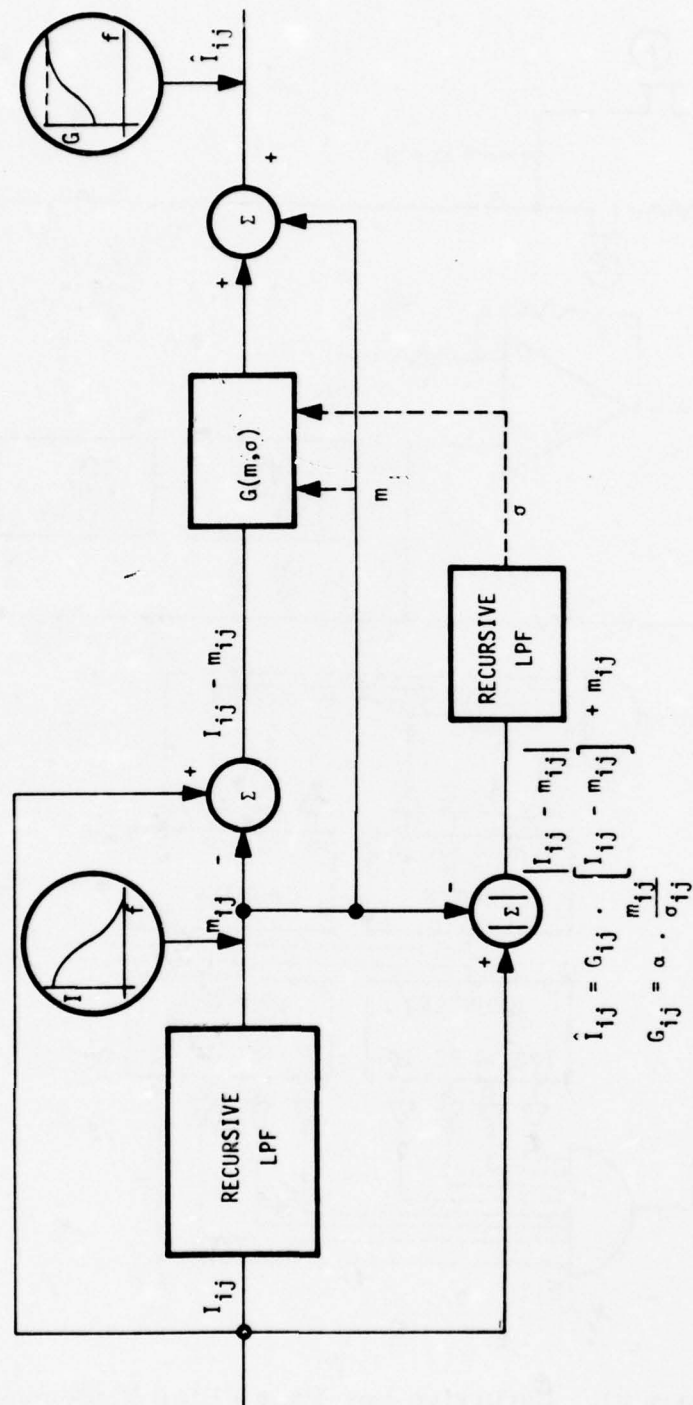
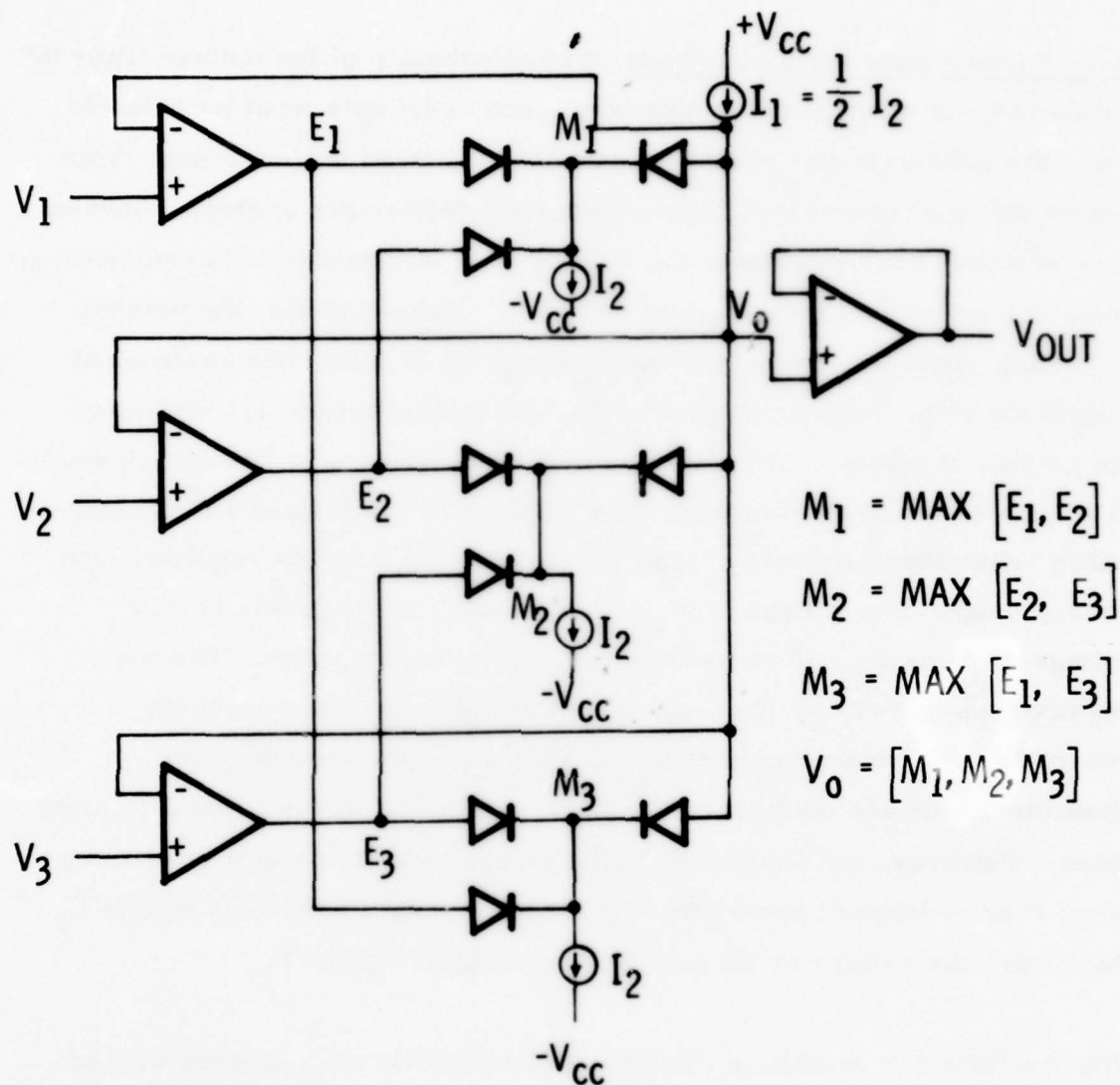


Figure 82. Recursive Approach to Local Area Gain/Brightness Control

MRT Enhancement Implementation--Implementation of the median filter is not as easy as most filtering functions since exact data must be retained, i. e., the median of five points is precisely the third largest point. This means that the transversal filters discussed earlier are useless. Several schemes have been proposed--for example, Westinghouse's thermometer-scale sorter-- to order groups of numbers. Unfortunately, the number of sorting operations required limits the speed to more than an order of magnitude below real time video rates. An analog circuit for obtaining the median of pixels, first proposed in 1967 in *Electronic Design*, is shown in Figure 83 for the three-pixel case. All of the amplifiers are in unity voltage gain configurations, since the back-to-back diodes result in zero voltage drop. The voltage at M_1 , for example, is the greater of the voltages E_1 and E_2 . If we assume both voltages are greater than the negative supply ($-V_{cc}$), then both diodes could conduct if connected separately. If we assume that $V_1 < V_2 < V_3$, then because diode characteristics are nonlinear, the diode associated with E_2 would conduct more. However, the diode associated with E_1 would now quit conducting since it is no longer biased properly. Hence the voltage at $M1$ equals E_2 . Similarly, the voltage at $M2$ equals E_3 and at $M3$ equals E_3 .

Now consider the remaining diodes. Since initially $-V_{cc}$ is greater than the voltages at $M1$, $M2$, or $M3$, all three diodes could conduct. However, since some diodes are reversed from before, the minimum voltage of $M1$, $M2$, or $M3$ will appear at V_0 . Thus V_2 becomes our output. The diode which is most conducting, naturally, is the one with greatest voltage drop across it; that occurs with minimum voltage at the diode cathode.



TO FIND MEDIAN OF N POINTS, WE NEED

$$N ! \left(\frac{N+5}{N+1} \right) \text{ DIODES}$$

Figure 83. Three-Pixel Median Filter

Unfortunately, however, the number of diodes increases very rapidly with increasing numbers of pixels. For a nine-pixel median we require 882 diodes, and for a 5 x 5 median filter we need more than 10^6 diodes. Simulations have shown, however, that a separable median filter in two dimensions, which uses two five-pixel median filters and four line delays, provides visual results virtually equivalent to the full 25-point median. The total number of diodes (50) required for a five-pixel median is easily integrated on a single chip. The unity-gain buffers can also be fabricated with either bipolar or MOS technology. However, standard MOS/CCD processing does not provide diode functions; only bipolar chips routinely include them. Unless we modify the MOS process to handle an extra diffusion, the best solution would be to fabricate the entire five-point median filter structure on a single bipolar chip, with current sources, diodes, resistors, and buffer amplifiers on the same chip. Then two such chips in conjunction with four line delays, could provide the full 5 x 5 separable median filter using a serial video data stream. The structure is illustrated in Figure 84.

Another technique for intraframe smoothing uses adaptive filters which are selected by calculation of the local curvature shown in Figure 85. We make local curvature calculations in each of the four directions shown and select the maximum value. If $X(i, j)$ is the input image and $Y_k(i, j)$ is the smoothed image as an output from the K^{th} filter, then the general form of the filter output is

$$Y_k(i, j) = \sum_{m=-3}^3 \sum_{n=-3}^3 e_k(m, n) x(i+m, j+n),$$

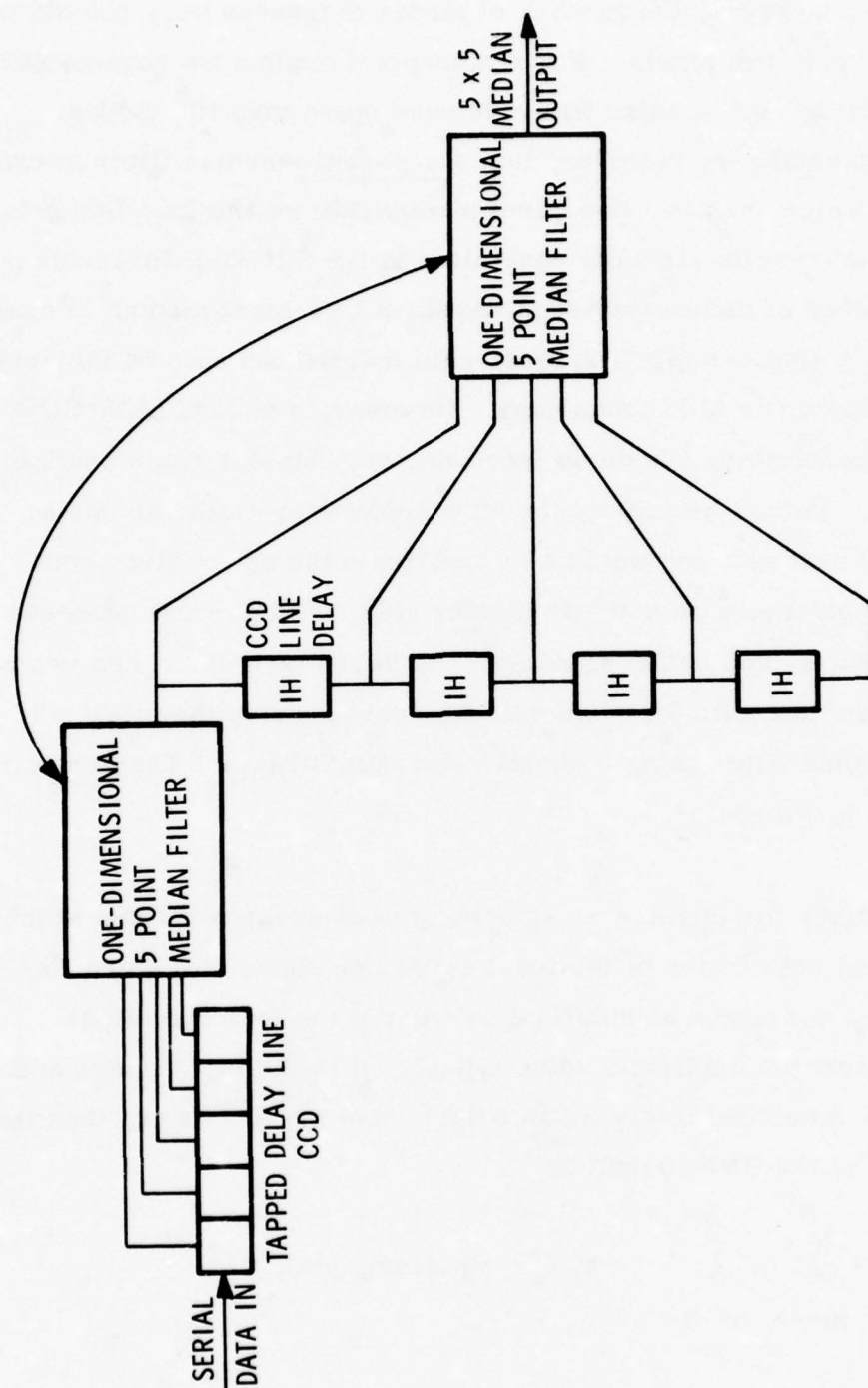


Figure 84. Separable Median Filter Structure

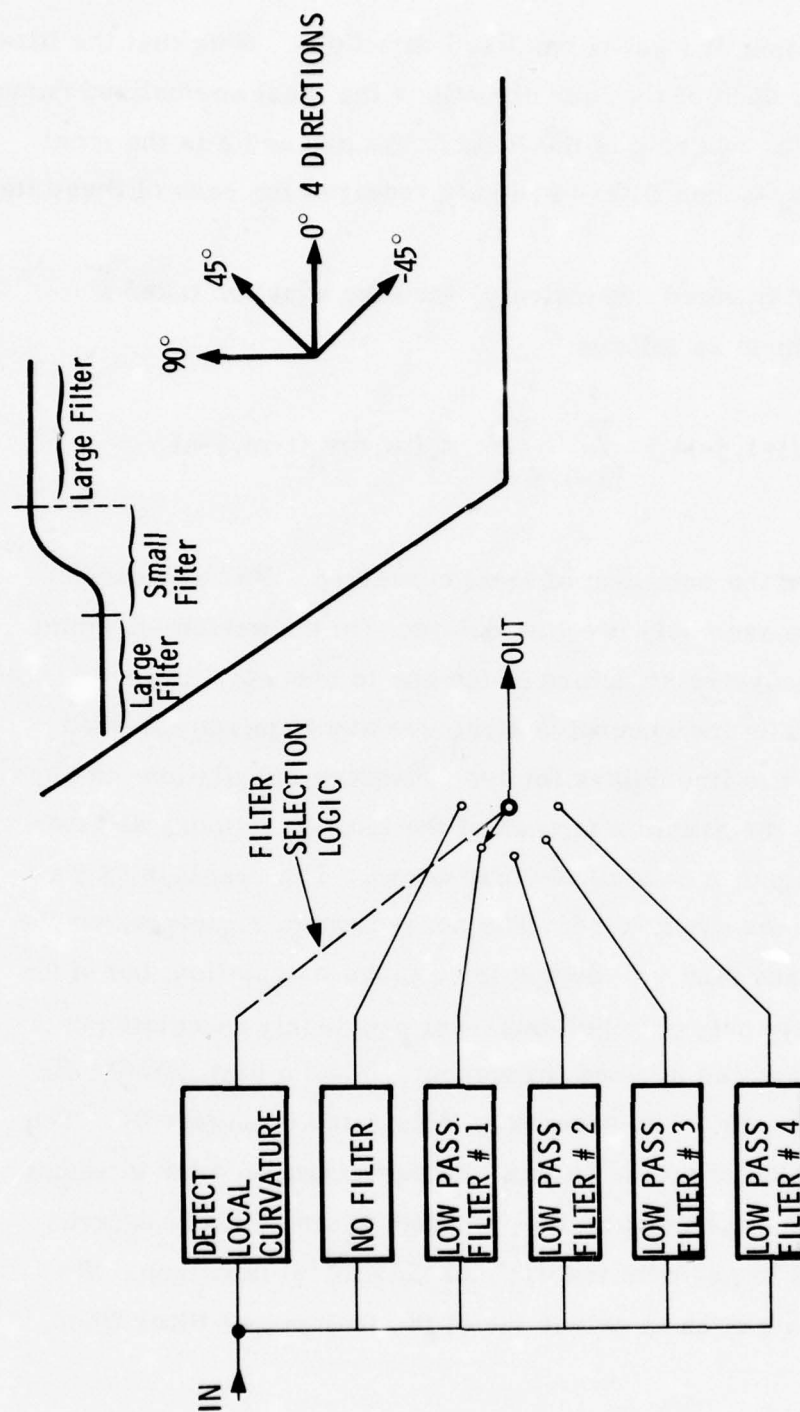


Figure 85. Adaptive Intraframe Smoothing

where e_k is a function of local normalized curvature. Note that the filter size is 7×7 . For each of the four directions the local normalized curvature is given by d/a , where d is the local curvature and a is the local gradient. Seven-by-seven filters are also required for each of these items.

If the filter is implemented recursively, then the equation takes a slightly different form as follows:

$$Y_k(i, j) = d_k x(i-1, j-1) + \sum_{m=0}^2 \sum_{n=0}^2 c_k(m, n) y_k(i-m, j-n)$$

Here c_k and d_k are the functions of local curvature. For a recursive implementation we need only two line delays. On the surface it would appear that the recursive structure is the one to choose. However, note that the line delays in the recursive structure are required for each calculation; i. e., two line delays for the curvature calculations in each direction, two for the gradient in each of the four directions, and two for the filtered output in each of the four cases. The result is that a total of 24 delay lines are needed. The nonrecursive structure, on the other hand, requires size line delays for a given calculation; but since all calculations rely only on pixel data--not previously calculated data--the same line delays can be used throughout. Thus a total of only six line delays is required. The scheme is illustrated in Figure 86. The CCD filters for calculating the output, the curvature, and the gradient are all in series and the outputs of y_1 through y_4 are merely delayed enough to allow us to perform the $d/1$ and $\text{MAX}(d/a)$ functions. We then merely gate a switch to select one of the four output filter lines.

The final MRT enhancement scheme we will discuss is frame registration. One scheme for providing purely translational registration uses the squared canonical production correlation to register on new frame $X_{k+1}(i, j)$ with another frame $X_k(i, j)$, as shown in Figure 87. We select a portion of the initial frame ($X_k(i, j)$) which has a bright edge--this could be done with one of the two-dimensional filters--and store that portion of the array (the $r(p, q)$). Before the region in the next frame which is centered about the reference region occurs in the serial data stream, we have had time to make the calculation of $f(r)$ ($r = 0.04\bar{r}$). The time available would be about 1/60 second, or one field time. Then we must calculate the quotient of two correlations: the first is a correlation of the function of reference points calculated above, and the second is a shifted autocorrelation of the new frame as a normalized function. The calculation of the function of $r(p, q)$ would be done digitally since so much time is available. (When compared to 10 to 20 MHz data rates, 1/60 second is an eternity.) The correlations, however, would be done with analog CCDs.

Figure 88 shows the hardware implementation. Digital counters are provided to identify the location of the bright edge (the reference window data). The A/D converter feeds all data to a microprocessor for storage and calculation. The microprocessor then gates the analog data from the next frame and also provides the $f(r)$ ($r = 0.04\bar{r}$) data for correlation. Two analog CCD correlations provide real time correlation, and the maximum quotient of the two indicates the correlation peak of the new frame.

Lest the analog correlators seem remote, we should point out that Honeywell has developed a 32-, 96-, and 128-point analog correlators which can operate at rates above 2 MHz. Redesign of input and output

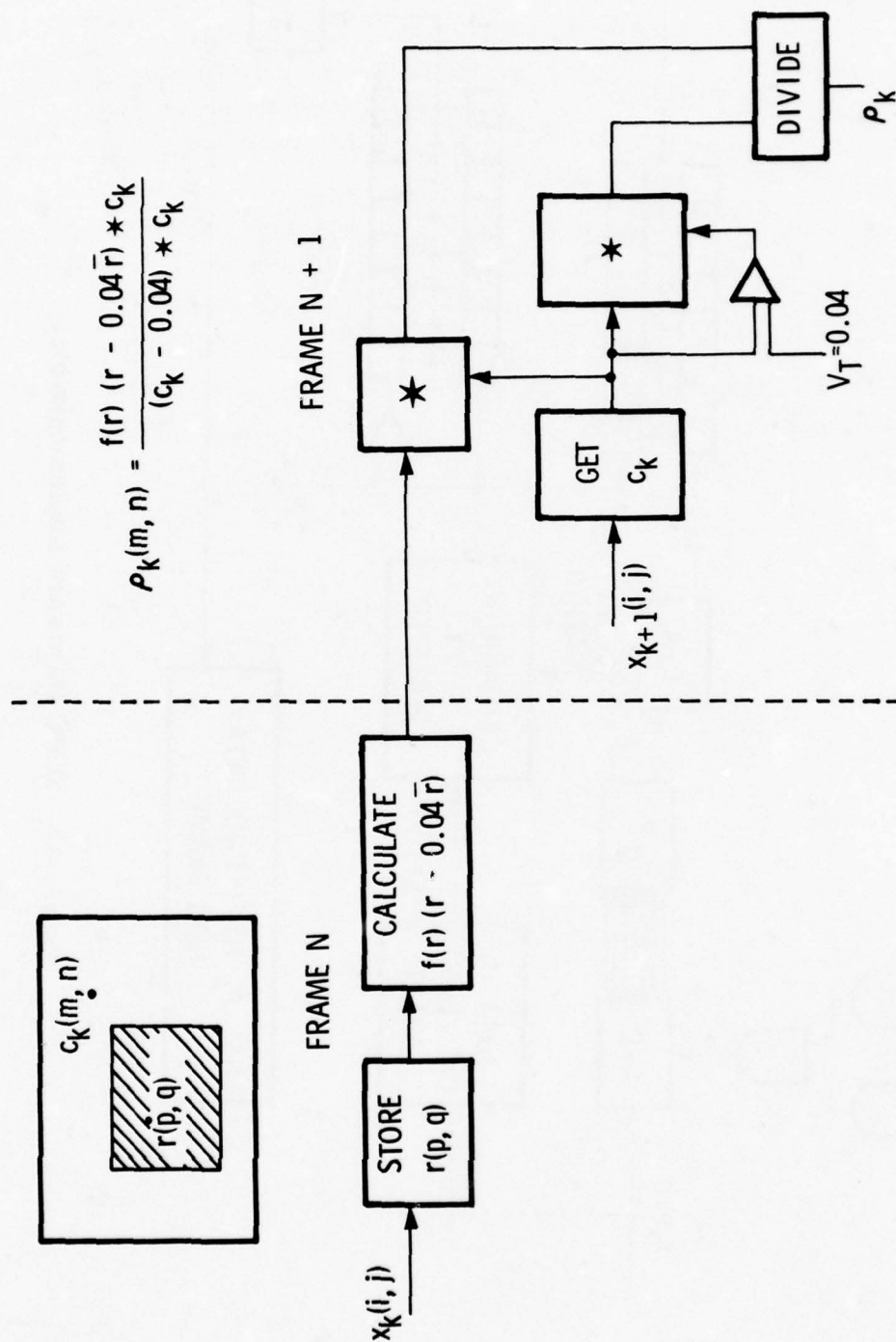


Figure 87. Squared Canonical Product Correlation for Frame Registration

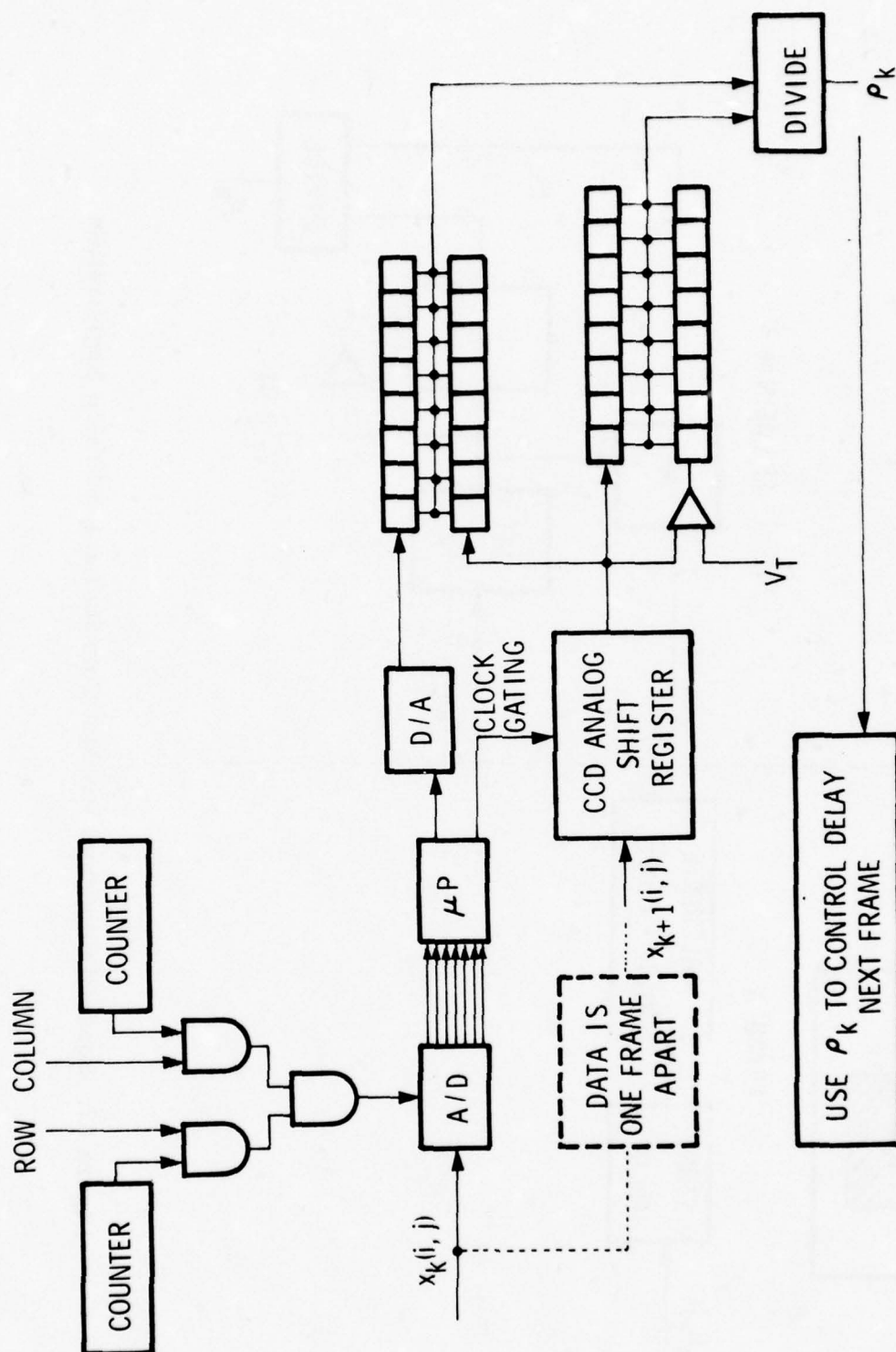


Figure 88. SCPC Hardware Implementation

circuits should increase the speed of these chips. Current dynamic ranges are in excess of 45 dB; linearity measurements are not available at this writing.

Resolution Restoration Implementation--In the final area we will discuss resolution restoration, both focus restoration and superresolution. One technique investigated was inverse Wiener filtering. As pointed out in the April 1977 interim report, the Wiener inverse filter shape closely approximates a high frequency emphasis filter. The two-dimensional transform of the true inverse filter is awkward to implement. In addition, exact knowledge of the point spread function (PSF) is difficult to acquire because the PSF is linear shift variant over area. For these reasons we recommend approximation of the Wiener filter with a simple high frequency emphasis filter. A recursive implementation of this filter was discussed earlier in this section.

The other algorithm for resolution restoration is the stochastic approximation algorithm. This algorithm uses an iterative approach to generate a superresolved image. The recursive equation used here is the following:

$$Y_{k+1}(i,j) = Y_k(i,j) + \mu_k [X_m(i,j) - h * Y_k(i,j)],$$

where $Y_k(i,j)$ is the k^{th} iteration of the output image, h is a system transfer function, $X_m(i,j)$ is a magnified input image, and μ_k is a function only of the iteration number.

We envision the resolution restoration function as being used over a small portion of the image. In addition, repeated iterations (approximately 10)

are required. Both of these factors will influence our implementation. The first step in generation of the superresolved image is generation of a magnified image as shown in Figure 89. Embedding zeroes can be done quite easily by using a series-parallel-series analog buffer for the input data. The input series register is clocked at twice the input data rate and the output series register is clocked at four times the input sampling rate. This mismatch is not as bad as it seems, at first glance; because of the repeated iterations the stochastic approximations algorithm is not geared to real time implementation.

The remainder of the stochastic approximations implementation is shown in Figure 90. A two-dimensional nonrecursive filter (approximately 7×7) is applied to the CCD analog memory output; this magnified and smoothed image is $X_m(i, j)$. The PROM shown provides the μ_k . A CCD fixed-weight filter performs the convolution of h with the previously iterated image. Note that the filter shows two apparent outputs. The first is the CCD filter output which is fed back to the differenced with the input magnified image. Note that the input sub-image is fed into the CCD analog memory only once; the contents are then read repeatedly. The other CCD filter "output" is the sequence of analog charge packets passing under the clock electrodes, i.e., the delayed filter input signal. A summing amplifier generates the analog output.

X	X	X	X	X
X	X	X	X	X
X	X	X	X	X
X	X	X	X	X
X	X	X	X	X

$$= X(i, j)$$

X		X		X		X		X
X		X		X		X		X
X		X		X		X		X
X		X		X		X		X
X		X		X		X		X

INTERPOLATE = $x_m(i, j)$ [illegible]

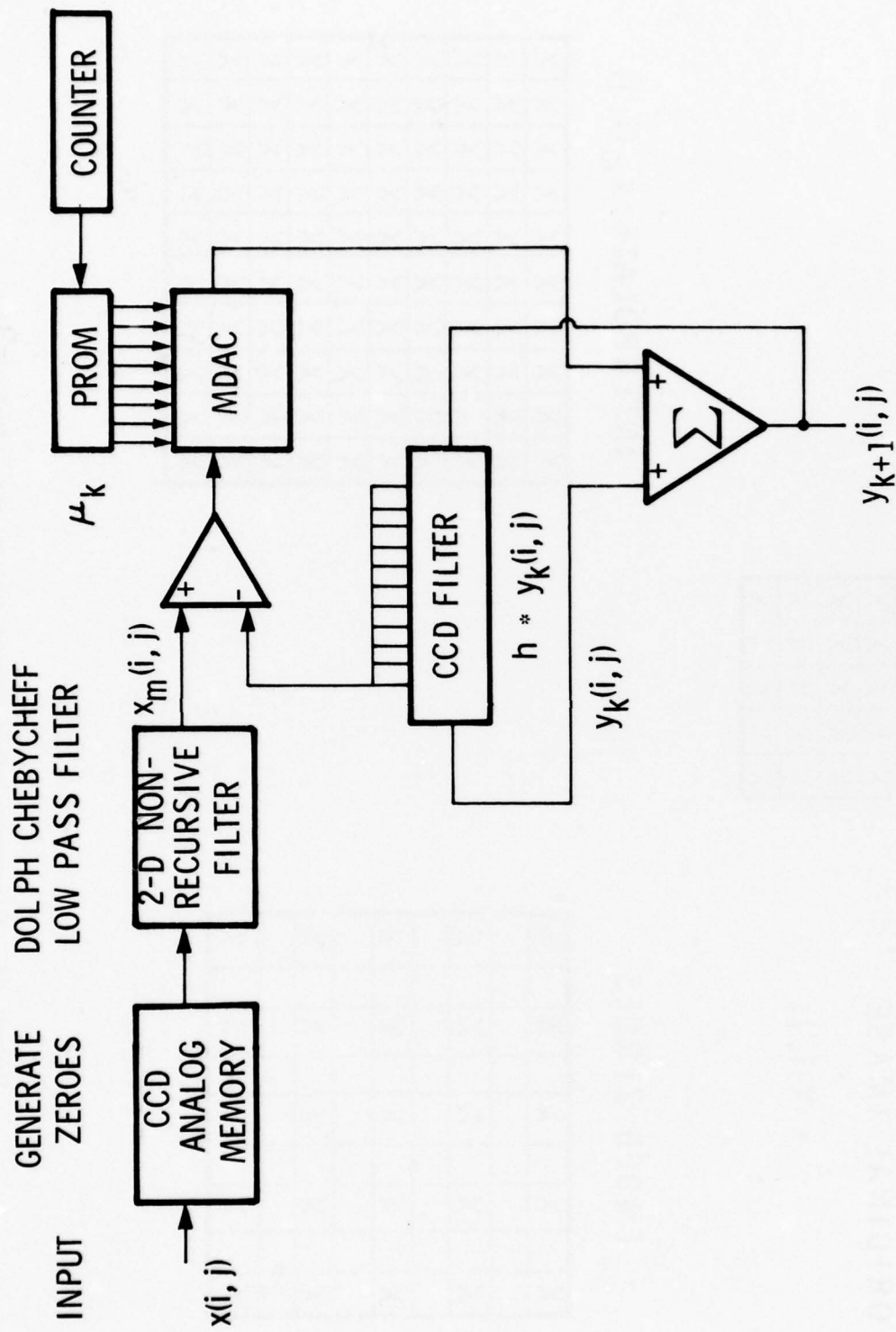


Figure 90. Stochastic Approximation Implementation

PART B
THE EFFECT OF CHROMIUM TUNGSTATE
ON A REDUCED RATE OF TUNG STATE
CHROMIUM TUNGSTATE ON THERMAL STABILITY

by
J.R. HARRISON
J.J. LINDEN
E.R. DUNN
J. WILK

ACKNOWLEDGEMENTS

We wish to thank Bud Hergenhahn and his staff in the Department of Psychology at Hamline University for their help and for allowing us to set up our apparatus in their facility. We also wish to thank Bryan Buck and Kurt Graffunder for constructing the experimental apparatus; and Doug Perlick for his help in reducing and analyzing the data.

SECTION I

INTRODUCTION

In conducting the various tasks that are part of airborne missions, pilots and navigators make use of radar, television and thermal sensor information. The quality of this information is dependent upon the parameters of the sensor and the characteristics of the display conveying it. Several ways of enhancing the quality of thermal images have been explored by Honeywell, working under contract number DAA953-76-C-0195 for the Night Vision Laboratory. As part of this work, we carried out an experimental evaluation. Its objective was to discover whether or not changes in image quality, caused by the various enhancement techniques, resulted in improvements in the performance of observers.

Sensor-display systems may deliver images of poorer quality than necessary because they have a limited range, add noise, and/or blur the image. Three methods of enhancement were considered in order to eliminate or reduce the effect of each of these problems:

1. Contrast Enhancement--When a wider range of temperatures exists in the original scene than can be accommodated by the sensor-display system, low contrast ground details may be lost. Contrast enhancement techniques may be used to accentuate these low contrast details (and to compress the overall temperature range of the scene).

2. Minimum Resolvable Temperature Enhancement--Some thermal images may be partially obscured by noise. By smoothing out noise, the minimum resolvable temperature may be lowered, so that some, previously invisible targets can be seen. There may be some adverse effects, since crucial details may be blurred by the smoothing process.
3. Resolution Restoration--The imagery may be blurred by the optics and electronics of the sensor. In this case, using information obtained by inspecting the degraded image, an attempt is made to model, and then correct for, the effects of blur, to produce a new, clearer image, which "restores" resolution.

Algorithms were developed to implement each of these techniques. The values of the parameters in the algorithms could be varied. In addition, since some images could have more than one of the problems outlined above, combinations of the three kinds of techniques were used.

A number of thermal images, all views of the ground and containing one or more military vehicles, were transformed using the various algorithms. The resultant images were available for this evaluation study. They were presented to observers, who were given the task of trying to decide whether particular hot spots were tanks, armored personnel carriers, trucks or jeeps. The speed and accuracy of the observers' responses were measured and used to determine the effectiveness of the enhancement algorithms.

SECTION II

METHOD

IMAGERY

There were forty test images and twelve additional, filler images. Both sets of images were taken with thermal sensors (either a thermoscope or a forward looking infrared sensor). They were ground views taken from the air, from elevated positions, or from the ground. Fifty-one of the images contained one, two or three military vehicles. The vehicles were some combination of tanks, jeeps, trucks and armored personnel carriers. The fifty-second image showed some cattle grazing. Both the test and filler images were taken from various viewing distances and aspect angles.

One vehicle was designated as the target for each image (one, large animal was selected on the image with cattle). Unfortunately, a disproportionately large number of vehicles were tanks. Because of this, on those images with more than one type of vehicle, tanks were not selected. Then, the number of test images with each type of target was: 23 - tanks, 5 - jeeps, 3 - trucks, 8 - armored personnel carriers, 1 - other.

Two strategies were adopted in order to reduce further the proportion of images with tanks as designated targets. The first involved those images with non-tank targets, that were enhanced by the resolution restoration method. Before resolution could be "restored", it was necessary to enlarge the image. This led to there being two treatment conditions with enlarged

images. To each sequence of images which were prepared for presentation, and in which there were original-sized images with non-tank targets, enlarged images with these same targets were added. Similarly, if the original sequence included enlarged non-tank target images, original-sized images were added. In addition to the added images being of different size, they were reversed left-to-right. The results of a brief pilot experiment had indicated that, under these conditions, observers were unlikely to notice that they had, in fact, seen two versions of the same basic image.

The second strategy was to incorporate a number of additional "filler" images into the sequence of test images. By using these two strategies, the proportion of images with tank targets seen by the observers was reduced considerably. Each group of observers was shown two sequences of 58 test slides. In each experimental session, the number of images with each type of target was: 46 - tanks, 21 - jeeps, 13 - trucks, 32 - armored personnel carriers, 4 - other.

Table 8 gives details of the forty test, and twelve filler, images. Those images with the prefix "H" in the first column were taken with a forward looking infrared sensor; the other images were taken with a thermoscope and provided by the Night Vision Laboratory. The second column indicates the number assigned to each image for this experiment. The remaining columns list the type of target designated, and whether any other vehicles are shown in the image.

TABLE 8. DETAILS OF IMAGES USED IN EXPERIMENT

NVL Designation	Test Number	Test Target	Other Vehicle(s)
1.1	11	Tank	--
1.3	1	APC	Tank
1.4	2	Jeep	Tank
1.5	12	Tank	--
1.6	3	Jeep	Tank, APC
1.8	13	APC	Tank
1.9	4	Jeep	Tank, APC
2.1	35	APC	Tank
2.6	36	APC	Tank, Jeep
2.16	37	APC	Tank
2.19	38	Jeep	Tank, APC
2.21	14	APC	Tank, Jeep
3.2	15	Tank	Tank
3.3	26	Tank	--
3.4	27	Tank	--
3.8	16	Tank	--
3.9	28	Tank	--
3.13	29	Tank	--
3.14	17	APC	--
3.15	30	Tank	--
3.17	31	Tank	--
3.18	32	Tank	--
4.1	18	Cow	Cows
4.3	19	Tank	--
4.6	5	Tank	--
4.8	39	Tank	--

TABLE 8. DETAILS OF IMAGES USED IN EXPERIMENT (concluded)

NVL Designation	Test Number	Test Target	Other Vehicle(s)
4.11	20	Tank	--
4.14	6	Tank	--
5.8	21	Tank	Tank
5.11	10	Tank	--
5.16	7	Tank	Tank
5.22	22	Tank	Tank
5.25	33	Tank	--
H.5	23	Tank	--
H.6	40	APC	--
H.10	24	Truck	--
H.12	25	Truck	--
H.15	8	Jeep	--
H.24	9	Truck	--
H.30	34	Tank	--

NVL Designation	Filler Number	Target	Other Vehicle(s)
2.3	45	Jeep	Tank, APC
2.5	41	APC	Tank
2.8	46	APC	Tank
2.17	42	APC	Tank
2.18	47	APC	Tank
2.23	48	Jeep	APC
2.24	49	APC	Jeep, Tank
2.25	50	Jeep	Tank, APC
2.26	43	APC	Tank
2.27	51	APC	Tank
3.6	52	APC	--
H.13	44	Truck	--

DESIGN AND EXPERIMENTAL CONDITIONS

The objective of this experiment was to discover whether observer performance was affected by changes in image quality caused by various enhancement techniques. Test images were treated by means of contrast enhancement, minimum resolvable temperature enhancement and resolution restoration algorithms, as well as by a cascade process that combined some of these algorithms. Since the original images appeared to be relatively noise free, and since minimum resolvable temperature enhancement techniques were developed to improve noisy thermal images, noise was added to a subset of the test images.

Seven different enhancement treatments and three combinations of treatments were used on the original images. With the subset of noise-added images, four enhancement treatments and three combinations were employed. Table 9 details the nineteen experimental conditions.

The various enhancement techniques were not employed systematically on all forty test images. The images were divided into five groups. Table 10 shows which enhancement conditions were used with each subset of images.

There were nineteen conditions to be tested. The minimum number of observations per image-by-treatment combination tested was set at ten. It was decided that it was not advisable to show any basic image, unless it was changed in size, more than two times to any observer. Thus, a minimum of 95 observers was required.

TABLE 9. EXPERIMENTAL CONDITIONS

Treatment	Original Scene	Treatment	Noise Added to Original Scene
1.	Original Scene	12.	Noise Added to Original Scene
2.	<u>A(a) Contrast Enhancement</u>		<u>B(a) Contrast Enhancement</u>
3.	High Frequency Emphasis Recursive		
4.	Local Area Gain Brightness Control Recursive	13.	Local Area Gain Brightness Control Recursive
	Local Area Gain Brightness Control Nonrecursive		
5.	<u>A(b) Minimum Resolvable Temperature</u>		<u>B(b) Minimum Resolvable Temperature</u>
	Recursive Adaptive Smoothing Filter	14.	Recursive Adaptive Smoothing Filter
6.	5 x 5 Median Adaptive Smoothing Filter	15.	Nonrecursive Adaptive Smoothing Filter
7.	<u>A(c) Resolution Restoration</u>	16.	5 x 5 Median Adaptive Smoothing Filter
8.	2 x 2 Digital Magnification		<u>B(c) Resolution Restoration</u>
	Stochastic Approximation Resolution Restoration		---
9.	<u>A(d) Cascade Process</u>		<u>B(d) Cascade Process</u>
10.	Combination of Treatments 3 and 5	17.	Combination of Treatments 13 and 14
11.	Combination of Treatments 3, 5 and 7	18.	2 x 2 Digital Magnification of Condition 17
	Combination of Treatments 3, 5 and 8	19.	Combination of Condition 17 and Stochastic Approximation Resolution Restoration

TABLE 10. TREATMENT CONDITIONS USED WITH EACH
SUBSET OF IMAGES

Treatment Conditions	Image Group	Image Test Numbers
1 to 12	I	1 to 10
12 to 19	II	1 to 9
1 to 8	III	11 to 25
1 to 6	IV	26 to 34
1, 7 and 8	V	35 to 40

Nineteen groups of observers were used, and nineteen different combinations of image and treatment conditions were prepared. The first group received combinations one and two. The second group received two and three. The third, three and four. And so on, until the nineteenth group, which received combinations nineteen and one. Therefore, each combination was presented to one group of observers first, and to another group second.

The combinations of image and treatment condition were prepared as follows. For images 1-9, which were used for all nineteen conditions, a different condition was selected for each image in each combination. For example, the first combination was as follows: image 4 - condition 1, image 2 - condition 3, image 3 - condition 5, image 4 - condition 7, image 5 - condition 9, image 6 - condition 11, image 7 - condition 13, image 8 - condition 15, image 9 - condition 17. For the second combination, images 1 and 2 were

paired with conditions 2 and 4, and the conditions were shifted similarly for the other images. This process continued throughout the nineteen sets of combinations, until each image had been paired with each condition.

Image 10 appears in Image Group I but was omitted from Group II, because it proved impossible to indicate, with any certainty, where its designated target was located when noise was added. This image was, therefore, used with only eleven conditions. One of these conditions was selected in a random fashion for each of the first eleven sets of image-condition combinations. In the remaining eight sets, eight of the eleven conditions were repeated.

The third image group, consisting of images 11-25, was used with eight conditions. These images were treated in a similar way to images 1-9. In the first combination set, images 11 and 19 were paired with condition 1, images 12 and 20 with condition 2, 13 and 21 with condition 3, 14 and 22 with condition 4, 15 and 23 with condition 5, 16 and 24 with condition 6, image 17 and 25 with condition 7, and image 18 with condition 8. Then in the second combination set images 11 and 19 were paired with condition 2. And so on until, after eight combination sets, each of the fifteen images in this Image Group had been used with each of the eight conditions. The cycle was repeated for combination sets nine through sixteen, and then begun again for the last three combination sets.

The fourth Image Group which included images 26 through 34, was used with conditions 1 to 6. A similar procedure to that adopted with Image Groups I and III was used, with the six conditions being cycled through

three times in combination sets one to eighteen, and with one more condition being repeated a fourth time for each image in the nineteenth set.

As described above, the forty test images were paired with a variety of enhancement conditions in each of the nineteen combination sets. For each one of these sets, the order of presentation of the test images, the twelve repeated, different-sized images, and six of the filler images, was randomized.

Each experimental session began with an experimenter describing the purpose of the study. Next, a set of instructions was read, then a series of four training slides were shown. This procedure gave the observers the opportunity to see what the various target vehicles look like when they are reproduced by a thermal sensor. Following this brief training, the first set of test images was presented. This took approximately forty minutes. There was a ten minute rest, then the second set of test images was given. The second set was shown with all its images reversed left to right. All the experimental sessions took between $1\frac{1}{2}$ and 2 hours to administer.

APPARATUS AND PROCEDURE

The experimental apparatus involved the following pieces of equipment:

- Screen
- Carousel projector
- Electrically operated shutter

- 14-channel tape recorder
- Power supply
- Six four-choice response boxes

Six observers could be tested at a time. There were three observer stations each side of a centrally placed projector. A shutter was placed in front of the projector. At the beginning of each trial, one experimenter announced the approximate location of the next target, then operated the shutter. The projected image fell on a two foot square screen area that was faintly marked with a six by six grid. If the designated target was still difficult to locate, a second experimenter indicated its position with a long pointer.

As soon as an observer recognized the target, he/she pressed the appropriate button on the four-choice response box in front of him/her. The amount of time elapsed since the shutter was opened and the type of response (i.e., tank, jeep, truck or armored personnel carrier) was recorded on the 14-channel tape recorder. After each test image had been shown for fifteen seconds, the first experimenter closed the shutter. If any observer had not yet responded, a "no" response was recorded with a time of fifteen seconds. The distance between the observers and the screen was approximately six feet.

OBSERVERS

The study was run in the Psychology Department of Hamline University. The observers were male and female students from various Psychology and Sociology courses. Data were collected from a total of 109 observers.

SECTION III

RESULTS

It was hypothesized that the accuracy of responses and the time taken to respond would be affected by the enhancement treatments. The Mann-Whitney two-sample, two-tailed U test was used to test these hypotheses. When the Mann-Whitney test can be compared with parametric tests, it is found to have a power-efficiency of 95 per cent (Mood, 1954).⁹ In addition, it can be applied directly to data which, like the response time data obtained here, are positively skewed and do not meet the assumptions necessary for parametric tests.

For testing purposes, comparisons were made between enhancement conditions and the appropriate unenhanced condition for each of the five Image Groups indicated previously in Table 3.

ACCURACY

The number of observers responding correctly, and the number of observers to whom each image-condition combination was presented, are shown in Table 11. The Table also shows the proportion of correct responses for each combination.

These proportions were used in running the Mann-Whitney test to compare the enhanced image conditions with the unenhanced conditions in each Image Group. For example, in Group I conditions 2 through 11 were compared

TABLE 11. NUMBER OF CORRECT RESPONSES (a), TOTAL NUMBER OF TRIALS (b), AND PROPORTION OF CORRECT RESPONSES (c) FOR EACH SCENE BY TREATMENT COMBINATION

		<u>Image Group I</u>											
		Treatment											
Scene		1	2	3	4	5	6	7	8	9	10	11	12
1	a	2	1	3	7	2	5	19	8	2	1	12	0
	b	11	12	11	22	10	11	38	22	16	21	37	11
	c	0.182	0.083	0.273	0.318	0.200	0.455	0.500	0.364	0.125	0.190	0.324	0.364
2	a	7	13	14	4	8	8	20	19	10	20	21	10
	b	13	23	23	12	11	10	31	22	12	10	35	10
	c	0.538	0.565	0.609	0.333	0.727	0.800	0.645	0.864	0.833	0.667	0.600	1.000
3	a	14	7	5	15	4	17	29	15	23	28	26	6
	b	15	12	10	21	11	23	33	20	24	32	36	11
	c	0.933	0.583	0.500	0.652	0.364	0.739	0.879	0.750	0.958	0.875	0.722	0.545
4	a	4	9	5	7	4	5	11	4	15	8	9	8
	b	11	14	9	12	13	16	33	22	22	31	22	11
	c	0.364	0.643	0.556	0.583	0.308	0.312	0.333	0.182	0.682	0.258	0.409	0.727
5	a	3	1	0	4	2	2	3	2	1	3	1	3
	b	11	11	11	14	15	12	13	12	6	12	11	9
	c	0.273	0.090	0.000	0.286	0.133	0.167	0.231	0.167	0.167	0.250	0.090	0.333
6	a	4	3	2	3	1	1	3	3	3	3	3	2
	b	10	11	11	11	11	14	15	12	13	9	11	12
	c	0.400	0.273	0.182	0.273	0.090	0.071	0.200	0.250	0.231	0.333	0.273	0.167
7	a	6	2	5	3	5	6	7	5	2	5	2	1
	b	11	10	10	11	11	11	11	14	15	12	10	5
	c	0.545	0.200	0.500	0.273	0.455	0.545	0.636	0.357	0.133	0.417	0.200	0.200
8	a	8	2	4	12	3	0	6	4	3	1	0	7
	b	26	9	11	20	15	11	35	33	11	23	27	23
	c	0.308	0.222	0.364	0.600	0.200	0.000	0.171	0.121	0.273	0.043	0.000	0.304
9	a	6	7	3	6	5	6	7	8	16	15	8	13
	b	10	10	12	12	11	10	31	37	22	26	11	24
	c	0.600	0.700	0.250	0.500	0.455	0.600	0.226	0.216	0.727	0.577	0.727	0.542
10	a	5	2	2	2	2	1	0	7	0	4	1	
	b	21	23	23	22	24	11	10	21	22	10	26	
	c	0.238	0.087	0.087	0.090	0.083	0.090	0.000	0.333	0.000	0.400	0.115	

TABLE 11. NUMBER OF CORRECT RESPONSES (a), TOTAL NUMBER OF TRIALS (b), AND PROPORTION OF CORRECT RESPONSES (c) FOR EACH SCENE BY TREATMENT COMBINATION (continued)

Image Group II								
Treatment								
Scene	12	13	14	15	16	17	18	19
1	a 4	3	2	3	3	2	6	19
	b 11	11	22	24	26	12	46	54
	c 0.364	0.273	0.090	0.125	0.115	0.167	0.130	0.352
2	a 10	16	17	14	18	18	19	7
	b 10	22	22	21	21	26	58	29
	c 0.000	0.727	0.773	0.667	0.857	0.692	0.328	0.241
3	a 6	17	16	20	22	10	38	27
	b 11	21	20	22	23	11	56	50
	c 0.545	0.810	0.800	0.909	0.957	0.909	0.679	0.540
4	a 8	12	8	8	9	11	9	12
	b 11	20	11	16	20	21	48	49
	c 0.727	0.600	0.727	0.500	0.450	0.524	0.187	0.245
5	a 1	1	2	0	0	1	3	7
	b 9	10	11	9	11	10	10	11
	c 0.111	0.100	0.182	0.000	0.000	0.100	0.300	0.636
6	a 2	2	2	0	2	4	4	2
	b 12	11	10	9	11	12	11	10
	c 0.167	0.182	0.200	0.000	0.182	0.333	0.364	0.200
7	a 1	1	5	2	0	2	2	4
	b 5	11	12	11	10	9	11	9
	c 0.200	0.090	0.417	0.182	0.000	0.222	0.182	0.444
8	a 7	1	3	11	2	4	5	3
	b 23	10	26	26	12	22	53	42
	c 0.304	0.100	0.115	0.423	0.167	0.182	0.094	0.071
9	a 13	6	15	8	0	22	29	1
	b 24	26	23	20	12	49	54	12
	c 0.542	0.231	0.652	0.400	0.000	0.449	0.537	0.083

TABLE 11. NUMBER OF CORRECT RESPONSES (a), TOTAL NUMBER OF TRIALS (b), AND PROPORTION OF CORRECT RESPONSES (c) FOR EACH SCENE BY TREATMENT COMBINATION (continued)

		Image Group III							
		Treatment							
Scene		1	2	3	4	5	6	7	8
11	a	13	12	13	10	5	5	11	13
	b	13	35	34	21	21	22	26	26
	c	0.394	0.343	0.382	0.476	0.238	0.227	0.577	0.500
12	a	14	17	21	24	10	10	12	20
	b	26	32	35	34	20	20	22	26
	c	0.518	0.515	0.600	0.706	0.500	0.500	0.545	0.769
13	a	7	8	9	12	7	6	46	42
	b	26	36	44	46	45	21	104	109
	c	0.269	0.222	0.205	0.261	0.156	0.286	0.442	0.385
14	a	4	9	4	8	7	4	34	22
	b	32	37	36	44	46	34	114	104
	c	0.125	0.243	0.111	0.182	0.152	0.118	0.298	0.212
15	a	19	12	11	20	23	27	26	14
	b	21	32	26	26	33	35	34	20
	c	0.905	0.545	0.538	0.769	0.697	0.771	0.765	0.700
16	a	3	1	13	8	2	5	7	12
	b	21	21	22	26	26	28	35	34
	c	0.143	0.048	0.591	0.308	0.076	0.179	0.200	0.353
17	a	25	13	12	23	16	8	57	51
	b	46	31	31	35	37	38	114	103
	c	0.543	0.419	0.387	0.657	0.432	0.211	0.500	0.495
18	a	0	0	0	0	0	0	0	0
	b	15	46	36	32	33	36	100	114
	c	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
19	a	9	5	3	2	3	5	7	6
	b	33	35	34	21	21	22	26	26
	c	0.273	0.143	0.088	0.095	0.143	0.227	0.269	0.231
20	a	3	3	6	7	3	3	2	3
	b	26	28	35	34	21	20	22	26
	c	0.115	0.107	0.171	0.206	0.143	0.150	0.090	0.115

TABLE 11. NUMBER OF CORRECT RESPONSES (a), TOTAL NUMBER OF TRIALS (b), AND PROPORTION OF CORRECT RESPONSES (c) FOR EACH SCENE BY TREATMENT COMBINATION (continued)

Image Group III (concluded)								
Scene	Treatment							
	1	2	3	4	5	6	7	8
21	a 1	1	2	9	8	5	2	3
	b 26	26	33	31	34	16	21	22
	c 0.038	0.038	0.061	0.290	0.235	0.312	0.095	0.136
22	a 11	18	17	17	13	5	15	13
	b 22	26	26	33	35	34	21	20
	c 0.500	0.692	0.654	0.515	0.371	0.147	0.714	0.650
23	a 1	6	6	6	6	5	6	2
	b 30	22	26	26	28	35	34	21
	c 0.033	0.273	0.231	0.231	0.214	0.143	0.176	0.095
24	a 14	17	13	7	15	5	42	33
	b 21	32	35	38	37	45	104	114
	c 0.667	0.531	0.371	0.184	0.405	0.111	0.404	0.289
25	a 7	8	7	9	17	6	29	17
	b 39	32	32	33	36	39	114	89
	c 0.179	0.250	0.219	0.273	0.472	0.154	0.254	0.191

Image Group IV						
26	a 28	16	24	13	17	10
	b 55	34	37	34	33	24
	c 0.509	0.471	0.649	0.382	0.515	0.417
27	a 12	31	23	26	13	18
	b 35	46	34	35	35	33
	c 0.343	0.674	0.676	0.743	0.371	0.545
28	a 19	14	21	22	18	9
	b 33	35	41	34	35	35
	c 0.576	0.400	0.512	0.647	0.514	0.257
29	a 26	23	21	42	28	19
	b 34	32	24	46	34	35
	c 0.765	0.719	0.875	0.913	0.824	0.543

TABLE 11. NUMBER OF CORRECT RESPONSES (a), TOTAL NUMBER OF TRIALS (b), AND PROPORTION OF CORRECT RESPONSES (c) FOR EACH SCENE BY TREATMENT COMBINATION (continued)

<u>Image Group IV (concluded)</u>								
Treatment								
Scene	1	2	3	4	5	6	7	8
30	a 24	26	23	26	21	14		
	b 35	34	32	35	46	34		
	c 0.686	0.765	0.719	0.741	0.674	0.412		
31	a 27	28	23	20	24	25		
	b 34	35	30	33	35	46		
	c 0.794	0.800	0.767	0.606	0.686	0.543		
32	a 28	20	26	19	12	19		
	b 46	34	35	30	33	35		
	c 0.609	0.588	0.743	0.633	0.364	0.543		
33	a 4	3	7	1	2	4		
	b 32	35	43	34	35	35		
	c 0.125	0.085	0.163	0.029	0.057	0.114		
34	a 25	21	29	32	23	23		
	b 35	32	35	46	34	35		
	c 0.714	0.656	0.829	0.696	0.676	0.657		
<u>Image Group V</u>								
	1						7	8
35	a 4						10	12
	b 27						67	66
	c 0.052						0.149	0.182
36	a 26						36	15
	b 70						81	67
	c 0.371						0.444	0.224
37	a 10						21	19
	b 67						70	81
	c 0.149						0.300	0.235
38	a 57						21	7
	b 76						57	70
	c 0.750						0.313	0.100

TABLE 11. NUMBER OF CORRECT RESPONSES (a), TOTAL NUMBER OF TRIALS (b), AND PROPORTION OF CORRECT RESPONSES (c) FOR EACH SCENE BY TREATMENT COMBINATION (concluded)

<u>Image Group V (concluded)</u>				
Treatment				
Scene	1		7	8
39	a 40		40	31
	b 70		80	67
	c 0 571		0 500 0 463	
40	a 34		47	53
	b 67		70	81
	c 0 507		0 671 0 654	

with condition 1. And in Group II, condition 1 (original scene) was compared with condition 12 (noise added); then condition 12 was used as the standard against which to compare conditions 13 through 19.

No significant differences were found for any of these comparisons. There were also no significant differences from Image Groups III, IV and V.

Table 12 summarizes the accuracy of response data for each Image Group. It shows the number and proportion of correct responses, incorrect responses and failures to respond for each condition.

RESPONSE TIME

The response times associated with the images in each Image Group were arranged in ascending order, for each condition. All failures to respond were included as 15-second response times. The Mann-Whitney test was then used to compare the enhanced conditions with the standard, unenhanced image conditions.

For Image Group I, condition 1 was compared with conditions 2 through 11. One of these comparisons involving condition 9 proved to be significant at a low significance level ($p < .093$).

In Image Group II, condition 1 was compared with condition 12, and then 12 was used as the standard of comparison for conditions 13 to 19. The comparison of 12 and 19 showed a significant difference at the $p < .037$ level.

For Group III, there were no significant differences in response time.

TABLE 12. TOTAL NUMBER OF TRIALS AND NUMBERS AND PROPORTIONS OF CORRECT RESPONSES, INCORRECT RESPONSES AND NONRESPONSES FOR EACH GROUP OF SCENES IN EACH TREATMENT TESTED

<u>Image Group I: Scenes 1-10</u>				
Treatment	Total Trials	Number Correct (Proportion)	Number Incorrect (Proportion)	Number Nonresponses (Proportion)
1	139	59 (.42)	61 (.44)	19 (.14)
2	135	47 (.35)	62 (.46)	26 (.19)
3	131	43 (.33)	64 (.49)	24 (.18)
4	159	63 (.40)	65 (.41)	31 (.19)
5	132	36 (.27)	76 (.58)	20 (.15)
6	129	51 (.40)	63 (.49)	15 (.12)
7	250	105 (.42)	111 (.44)	34 (.14)
8	215	75 (.35)	107 (.50)	33 (.15)
9	163	75 (.46)	70 (.43)	18 (.11)
10	206	88 (.43)	88 (.43)	30 (.15)
11	226	85 (.38)	108 (.48)	33 (.15)
12	116	54 (.47)	45 (.39)	17 (.15)

<u>Image Group II: Scenes 1-9</u>				
12	116	54 (.47)	45 (.39)	17 (.15)
13	146	59 (.40)	63 (.43)	24 (.16)
14	157	70 (.45)	62 (.39)	25 (.16)
15	158	66 (.42)	60 (.38)	32 (.20)
16	146	56 (.38)	65 (.45)	25 (.17)
17	172	121 (.51)	79 (.33)	23 (.13)
18	347	107 (.27)	225 (.57)	57 (.16)
19	266	62 (.22)	177 (.62)	55 (.21)

TABLE 12. TOTAL NUMBER OF TRIALS AND NUMBERS AND PROPORTIONS OF CORRECT RESPONSES, INCORRECT RESPONSES AND NONRESPONSES FOR EACH GROUP OF SCENES IN EACH TREATMENT TESTED (concluded)

Image Group III: Scenes 11-25

1	437	131 (.30)	252 (.58)	54 (.12)
2	462	130 (.28)	283 (.61)	49 (.11)
3	485	140 (.29)	281 (.58)	64 (.13)
4	480	162 (.34)	250 (.52)	68 (.14)
5	473	135 (.29)	270 (.57)	68 (.14)
6	445	99 (.22)	266 (.60)	80 (.18)
7	891	300 (.34)	483 (.54)	108 (.12)
8	854	251 (.29)	503 (.59)	100 (.12)

Image Group IV: Scenes 26-34

Treatment	Total Trials	Number Correct (Proportion)	Number Incorrect (Proportion)	Number Nonresponses (Proportion)
1	339	193 (.57)	127 (.37)	19 (.06)
2	317	182 (.57)	113 (.36)	22 (.07)
3	311	197 (.63)	92 (.30)	22 (.07)
4	327	201 (.61)	104 (.32)	22 (.07)
5	320	151 (.53)	116 (.41)	21 (.07)
6	312	150 (.48)	139 (.45)	28 (.09)

Image Group V: Scenes 35-40

1	427	171 (.40)	206 (.48)	50 (.12)
7	435	175 (.40)	228 (.52)	32 (.07)
8	432	137 (.32)	255 (.59)	40 (.09)

In Group IV, where conditions 2 through 6 were compared with condition 1, statistically significant differences were indicated for conditions 5 ($p < .030$) and 6 ($p < .00006$).

There were no differences that were significant for Group V.

Tables 13 and 14 summarize the response time data. Table 6 shows the geometric mean response time for each image-condition combination. Table 7 gives three measures of central tendency for each condition in each of the five Image Groups. The measures are the geometric mean, arithmetic mean and median response time.

The geometric mean ($= \sqrt[n]{x_1 x_2 x_3 \dots x_n}$) has been used recently by several investigators (e.g., Monk, 1974; Monk and Brown, 1975; Bloomfield, Graf and Graffunder, 1975)¹⁰⁻¹² for time data that are positively skewed. It is the exponential of the mean of the logarithmically transformed data, and appears to be the best available measure of central tendency for this type of data.

TABLE 13. GEOMETRIC MEAN TIMES FOR SCENE BY
TREATMENT COMBINATIONS, IN SECONDS
(The Geometric Mean is $\sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \dots \cdot x_n}$.)

Image Group I: Scenes 1-10 by Treatments 1-12

Scene	Treatment											
	1	2	3	4	5	6	7	8	9	10	11	12
1	8.27	7.44	7.10	9.81	7.33	5.30	5.13	4.53	6.13	3.90	5.29	4.57
2	6.27	7.21	5.62	6.94	4.06	6.53	3.06	3.11	3.56	5.06	4.59	2.62
3	4.13	4.34	7.15	4.61	6.81	4.34	3.86	5.34	4.38	2.67	4.56	4.49
4	5.07	4.59	7.73	5.64	7.76	4.81	5.56	6.18	4.02	5.67	5.01	4.50
5	4.36	4.38	5.79	7.35	6.78	6.78	8.50	8.46	13.00	6.46	6.50	8.08
6	8.75	6.67	7.97	4.62	4.84	7.55	6.36	7.09	7.74	11.17	10.76	8.24
7	6.37	4.90	7.83	5.28	5.78	4.53	4.70	5.89	8.11	8.60	8.70	12.53
8	7.29	5.88	5.58	4.39	5.59	6.49	7.32	7.74	5.30	7.76	7.74	6.71
9	4.34	3.83	7.14	6.79	5.48	5.56	7.73	5.94	4.63	5.97	4.69	5.62
10	4.32	5.32	4.16	6.43	6.85	3.95	3.92	7.67	3.66	6.95	7.63	*

* No data were collected for Scene 10, Treatment 12.

Image Group II: Scenes 1-9 by Treatments 12-19

Scene	Treatment							
	12	13	14	15	16	17	18	19
1	4.57	4.38	3.64	6.32	4.66	7.35	7.00	6.54
2	2.62	5.19	4.80	3.32	3.94	5.24	7.58	7.74
3	4.49	4.41	4.53	2.93	2.38	3.12	5.57	7.87
4	4.50	5.74	3.36	5.74	4.65	4.34	5.84	6.87
5	8.08	8.58	4.29	4.23	2.89	5.50	4.28	4.10
6	8.24	10.44	12.54	13.23	12.66	7.93	7.63	9.86
7	12.53	12.38	5.86	7.72	10.73	8.70	11.34	10.11
8	6.71	9.38	9.24	8.11	7.54	6.69	7.43	6.60
9	5.62	8.24	6.24	7.53	11.06	5.99	6.54	8.11

Image Group III: Scenes 11-25 by Treatments 1-8

Scene	Treatment							
	1	2	3	4	5	6	7	8
11	5.45	4.70	4.66	5.66	3.45	4.46	3.64	3.09
12	5.16	4.62	4.69	3.98	5.67	3.93	3.29	2.91
13	6.48	6.21	6.08	6.07	6.86	5.98	6.82	6.62
14	7.05	6.17	7.19	7.50	7.04	6.44	7.49	7.51

TABLE 13. GEOMETRIC MEAN TIMES FOR SCENE BY
TREATMENT COMBINATIONS, IN SECONDS
(The Geometric Mean is $\sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \dots \cdot x_n}$.)
(concluded)

Image Group III: Continued

Scene	Treatment							
	1	2	3	4	5	6	7	8
15	3.68	5.82	6.33	4.21	4.22	4.72	3.84	5.86
16	6.21	5.33	4.52	5.53	8.62	7.09	7.07	5.94
17	6.62	4.94	4.52	5.24	7.50	7.91	6.29	5.69
18	7.75	6.45	8.04	7.59	4.54	5.87	5.52	6.16
19	4.70	4.90	4.29	5.30	4.36	3.33	3.60	4.46
20	5.68	7.58	7.38	7.11	5.95	7.61	5.20	6.83
21	6.34	6.29	7.24	6.56	7.72	6.74	5.67	5.04
22	6.86	4.94	6.36	6.72	7.42	7.39	4.35	3.02
23	4.97	5.83	5.04	6.49	6.73	7.61	6.03	6.17
24	7.33	4.88	6.70	7.25	6.18	10.61	6.84	7.92
25	5.80	6.74	5.86	5.32	4.53	6.41	6.03	5.24

Image Group III: Scenes 26-34 by Treatments 1-6

Scene	Treatment					
	1	2	3	4	5	6
26	4.25	4.96	3.81	6.92	4.77	5.91
27	5.36	5.11	3.62	4.23	6.92	6.50
28	5.93	6.55	5.78	6.59	6.51	8.58
29	3.98	3.61	4.16	3.15	3.72	5.92
30	3.27	3.28	4.24	3.85	3.53	3.89
31	3.35	3.79	3.76	3.91	4.00	4.13
32	3.60	3.44	3.25	4.02	4.39	3.45
33	5.88	7.36	9.84	6.22	7.74	9.24
34	4.57	4.24	3.86	4.33	4.00	4.37

Image Group V: Scenes 35-40 by Treatments 1, 7 and 8

Scene	1	7	8
35	5.81	5.07	4.94
36	5.08	4.96	4.12
37	6.11	5.87	6.76
38	3.32	4.85	5.14
39	4.19	5.46	5.26
40	4.02	3.42	3.62

TABLE 14. ARITHMETIC MEAN, GEOMETRIC MEAN, AND MEDIAN TIMES FOR EACH GROUP OF SCENES IN EACH TREATMENT TESTED, IN SECONDS (The Geometric Mean is $\sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \dots \cdot x_n}$.)

Image Group I: Scenes 1-10

<u>Treatment</u>	<u>Arithmetic Mean</u>	<u>Geometric Mean</u>	<u>Median</u>
1	6.81	5.71	5.40
2	6.77	5.47	5.10
3	7.34	6.12	6.10
4	7.48	6.04	6.30
5	7.16	6.12	5.97
6	6.53	5.37	5.95
7	6.58	5.31	5.80
8	7.18	5.91	6.25
9	6.15	5.06	4.90
10	6.76	5.34	5.85
11	7.00	5.83	6.10
12	7.19	5.67	6.40

Image Group II: Scenes 1-9

12	7.19	5.67	6.40
13	7.98	6.70	7.20
14	6.85	5.54	5.50
15	7.20	5.68	5.55
16	6.67	5.15	5.20
17	6.94	5.76	6.05
18	7.79	6.68	6.80
19	8.20	7.11	6.87

TABLE 14. ARITHMETIC MEAN, GEOMETRIC MEAN, AND MEDIAN TIMES FOR EACH GROUP OF SCENES IN EACH TREATMENT TESTED, IN SECONDS (The Geometric Mean is $\sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \dots \cdot x_n}$) (concluded)

Image Group III: Scenes 11-25

1	7.07	5.95	6.25
2	6.76	5.64	5.60
3	6.98	5.86	6.00
4	7.06	6.02	6.07
5	7.14	6.01	6.15
6	7.62	6.43	6.55
7	7.00	5.93	6.05
8	7.05	5.95	6.05

Image Group IV: Scenes 26-34

<u>Treatment</u>	<u>Arithmetic Mean</u>	<u>Geometric Mean</u>	<u>Median</u>
1	5.28	4.31	4.30
2	5.69	4.57	4.65
3	5.75	4.56	4.80
4	5.54	4.57	4.60
5	5.87	4.82	5.30
6	6.62	5.40	5.55

Image Group V: Scenes 35-40

1	5.86	4.64	4.87
7	5.79	4.88	4.85
8	5.83	4.88	4.30

SECTION IV

DISCUSSION

In terms of accuracy of response, there were no statistically reliable differences between the unenhanced and enhanced conditions.

For the response time data, there were four statistically significant differences. The first of these was in Image Group I. Here, cascade condition 9 (which involved the combination of the local area gain brightness recursive contrast enhancement algorithm, with the recursive adaptive smoothing filter minimum resolvable temperature enhancement algorithm) resulted in a reduced response time. The geometric mean time was lowered from 5.71 to 5.06 seconds. The difference in the two distributions of times was significant at the $p < .093$ level.

In Image Group II, another cascade condition was significantly different from the unenhanced condition. However, this time there was an increase in response time (the geometric mean went from 5.67 to 7.11 seconds, $p < .037$). This enhancement condition (19) included a resolution restoration algorithm along with the same contrast and minimum resolvable temperature enhancement algorithms mentioned in the previous paragraph. These enhancement algorithms were used to treat the noise-added original images (condition 12).

The significant differences found for Group IV were also associated with decrements in performance. Both minimum resolvable enhancement conditions--5 (Recursive Adaptive Smoothing Filter) and 6 (5 by 5 Median

Adaptive Smoothing Filter)--resulted in increases in response time. For condition 5, there was an increase in geometric mean response time from 4.31 to 4.82 seconds ($p < .030$). With condition 6, the increase was from 4.31 to 5.40 seconds ($p < .00006$). Most of the images in Group IV have a single, very large target. It appears that the smoothing filters, instead of removing noise, in fact removed the internal detail of the targets so that more time was required by the observers before they could decide which target they were looking at. Interestingly, there was no statistically significant decrement in the accuracy of the responses for these two conditions.

CONCLUSIONS

We can summarize our findings as follows:

- The changes in image quality that occurred, when enhancement techniques were employed, did not lead to changes in the accuracy of recognition.
- A small improvement in response time occurred for noise free images when contrast and minimum resolvable temperature algorithms were combined.
- A combination of contrast, minimum resolvable temperature and resolution restoration algorithms produced longer response times with noisy images.
- The use of minimum resolvable temperature algorithms with large target images led to increases in response time.

COMMENTS

The conditions involving local area brightness control contrast enhancements might usefully be compared with data reported recently by Scanlon, Hersberger, and Herman (1977).¹³ They report no overall change in response time when an enhanced condition of this kind is compared with an unenhanced one. However, they suggest that if scene complexity is taken into account, there may be changes: with low complexity the enhancement may help, while with high complexity it may prove detrimental.

In the experiment reported here, no assessment was made of the relative complexity of the test images. If such information had been available, it would have added greatly to the generality of our results.

PART II
SEARCH EFFECT
MODEL DEVELOP

SECTION I

INTRODUCTION

OBJECTIVES

The objective of this program is to develop a model of target acquisition for displays from electrooptical sensors in tactical situations. A key requirement is to incorporate into the model the effects of variations in the target and background characteristics.

Assume that a sensor points somewhere in the direction of a military threat. What is the probability that an operator looking at the display will be able to see the target within a given interval of time? Anyone contemplating the use of a night vision device needs the answer to this question for any condition at hand. But, despite the volume of effort attacking this problem, an answer to the question is currently unavailable.

BACKGROUND

It is usually assumed that visual search performance is a function of such factors as target size, target-background contrast, image quality, the area to be searched, and time available. It is also well-known that target and background characteristics have an important effect on performance. However, search modelers typically take a very simple view of how to treat these effects. Often the relationship between the target

AD-A049 564

HONEYWELL INC MINNEAPOLIS MINN SYSTEMS AND RESEARCH --ETC F/G 17/5
AUTOMATED IMAGE ENHANCEMENT TECHNIQUES FOR SECOND GENERATION FL--ETC(U)
DEC 77

DAA653-76-C-0195

NL

UNCLASSIFIED

77SRC93

4 OF 4

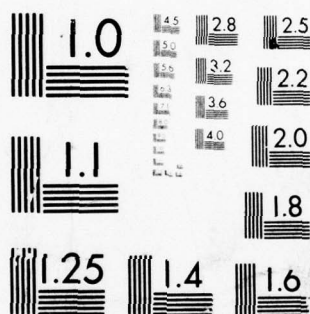
AD
A049564



END
DATE
FILMED

3 - 78

DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

SECTION I

INTRODUCTION

OBJECTIVES

The objective of this program is to develop a model of target acquisition for displays from electrooptical sensors in tactical situations. A key requirement is to incorporate into the model the effects of variations in the target and background characteristics.

Assume that a sensor points somewhere in the direction of a military threat. What is the probability that an operator looking at the display will be able to see the target within a given interval of time? Anyone contemplating the use of a night vision device needs the answer to this question for any condition at hand. But, despite the volume of effort attacking this problem, an answer to the question is currently unavailable.

BACKGROUND

It is usually assumed that visual search performance is a function of such factors as target size, target-background contrast, image quality, the area to be searched, and time available. It is also well-known that target and background characteristics have an important effect on performance. However, search modelers typically take a very simple view of how to treat these effects. Often the relationship between the target

and background is encompassed by a single parameter that allows one to fit the data to the model but that has no predictive power. On occasion an analysis has been made of the factors which make up (i. e., predict to) complexity or difficulty.¹⁴ There has also been considerable effort to develop methods of defining and measuring background texture as exemplified by the work of Rosenfeld and his coworkers.¹⁵

THE SEARCH MODEL

The immediate objective was to begin the development of a search model that fully incorporates the effects due to characteristics of the target and the background--size, shape, internal details, texture, and so on.

Someday we may have a model that predicts to all situations we can contemplate, but for now we must limit ourselves in a number of ways. First, the situation is static. Over a period of time there is no change in the search field--no perspective change, no target movement, no distance change. Second, there is no potential for changing the field; the target is always in the field or, for some purposes, the target is out of view and will never come into view.

The target will be relatively free from contact with other objects. It will not be occluded, nor will it abut other things in the field. Another limitation is that the fields are achromatic with no immediate plans for extending to colored fields. There are presently other restrictions. It is assumed that there is no pre-briefing. Each search field is

assumed to contain no more than one target, and the targets are assumed to belong to a limited class, namely small military objects (trucks, guns, vans, etc.).

The general nature of search is well understood. It has often been described as a combination of central and peripheral visual processes. An object is detected extra-foveally; the gaze is directed toward that object, and it is examined foveally. Thus, to construct a model we must predict extra-foveal conspicuity and foveal recognizability. However, there is another matter that is often recognized but universally ignored in formal search models. That is, the search field is divided into different, more or less homogeneous, regions with the conspicuity of the target depending on the region. Once the observer views the scene, each region is assigned a probability of containing the target. The observer then decides which regions to search first and how long to stay within each region.

Thus, when searching a given scene, there are three primary processes:

1. Strategy in selecting a region to search and remaining within it;
2. Extra-foveal detection and discrimination within a region; and
3. Foveal recognition.

A search model which encompasses the effects of target and background characteristics must include all three processes.

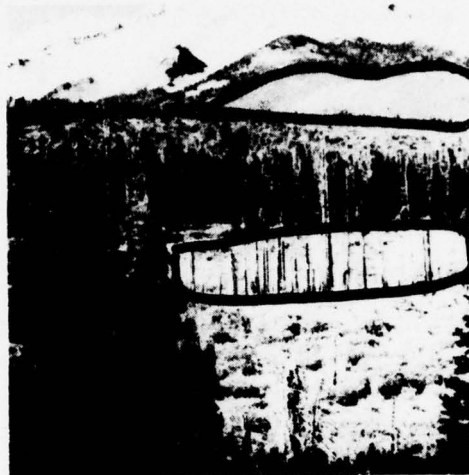
During the present study we observed the search for targets in homogeneous regions. By doing this we were able to remove selection strategy as a significant factor. Thus, we were able to isolate the effects of target and background characteristics. We feel that we must be able to predict for this simpler situation if we are to have any hope at all of predicting to the more general one. Examples are shown in Figure 91.

The problem then is to determine, for a range of homogeneous backgrounds, which measurable characteristics of targets and backgrounds are related to performance.

THE MATHEMATICAL STRUCTURE OF SEARCH FOR A TARGET IN A HOMOGENEOUS FIELD - INITIAL FORMULATION

We first attempted to develop a new statistical mathematical model for search based upon continuous Markov processes. In particular, the model for Brownian motion was examined. Notwithstanding the difficulties associated with the solution of complicated stochastic differential equations, this approach was found deficient for two essential reasons:

1. Visual search is not really random in the sense of a random walk or Brownian motion. If we assume a random walk, we come up against experimental data that clearly contradict the basic assumptions (e.g., equiprobable "jump" states).
2. This cannot be repaired by assuming a "force field." We would then encounter velocity and acceleration effects in contradiction to experimental evidence, which shows a more or less constant fixation rate.



1

2



3

4

Figure 91. Homogeneous Regions

However, an extensive survey of the literature on search models led us to conclude that a simple and straightforward approach would be adequate for our purposes. That is, the probability of finding the target by time t is given by

$$P(t) = K(1 - e^{-\lambda t}) \quad (14)$$

where λ and K are both functions of the scene parameters. This was concluded by noting that almost all the models surveyed reduced essentially to this form.

For example, the General Research Corporation (GRC) Model A is given by (with some simplification)

$$P(t) = P_1 P_2 P_3 P_4$$

where

$$P_1 = 1 - (1 - P_g)^n$$

P_g is the single glimpse probability of detection and n is the number of glimpses in a single time frame, where

$$P_2 = \begin{cases} 1 - e^{-((N-3.2)/11)^2} & \text{for } N \geq 3.2 \\ 0 & \text{for } N < 3.2 \end{cases}$$

and

$$P_3 = \frac{1}{1 + \left(\frac{M}{29T^{0.93}} \right)^{1.29}}$$

and

$$P_4 = \begin{cases} 1 - e^{-(S/N-1)} & \text{for } S/N \geq 1 \\ 0 & \text{for } S/N < 1 \end{cases}$$

Ignoring, for this purpose, the definitions of the parameter N, M, T, S, we note that we can rewrite these equations as

$$P_1 = 1 - e^{n \ln(1-P_g)}$$

Or, setting $nT = t$, we have

$$P_1 = 1 - e^{t[1/T \ln(1-P_g)]}$$

or

$$P_1 = 1 - e^{-\lambda t}, \quad \lambda = -\frac{1}{T} \ln(1-P_g)$$

and $K = P_2 P_3 P_4$ is a function of N, M, T and S. Thus, the GRC model reduces to the standard form.

To use this form of $P(t)$ we must determine, for each experimental parameter set (each scene), values for K and λ . Thus,

$$K = F(p) \text{ where } p = (p_1, p_2, \dots, p_m)$$

$$\lambda = G(q) \text{ where } q = (q_1, q_2, \dots, q_n)$$

and where $\{p_i\} \cap \{q_j\}$ is not necessarily empty. In general these systems will be nonlinear and overdetermined, so we cannot solve them in closed form directly.

PRESENT FORMULATION

A graph of Equation (14) is shown in Figure 92.

The two parameters in the equation have a direct interpretation. K can be interpreted as a measure of the recognizability of the target; given enough time $P(t)$ approaches K . When there is sufficient detail, $K = 1$, and $P(t)$ approaches 1. λ is the difficulty of detecting or locating the target. However, the initial examination of the data from the experiment showed Equation (14) to be inadequate.

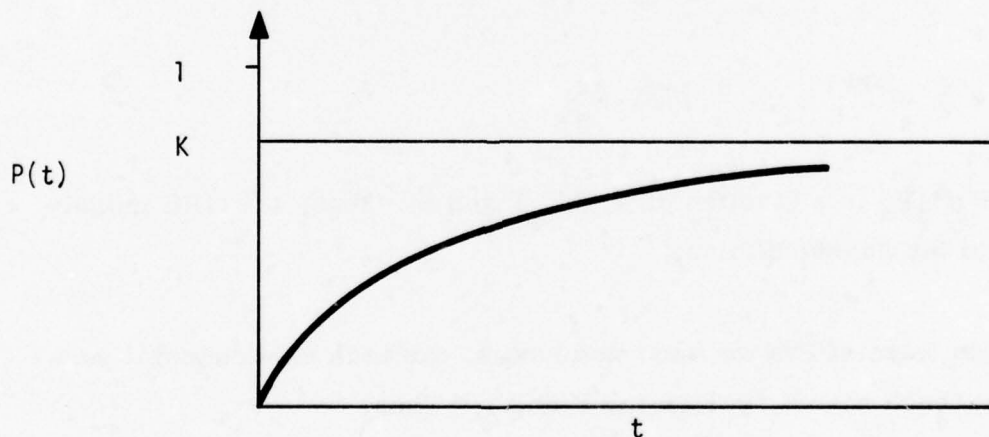


Figure 92. Graph of Equation (14)

In fact, the cumulative probability data have the general form shown in Figure 93 and are expressed by Equation (15).

$$P(t) = \begin{cases} K(1 - e^{-\lambda(t-d)}) & \text{for } t \geq d \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

$P(t)$ does not rise above 0 until time d . Thus, d represents a delay which relates to two main factors. First, there is an initial orientation time, when the entire scene is being inspected. Second, there is a response time, which includes both a recognition time as well as a time to respond to the requirements in the experimental situation. The reason experimental data do not usually appear to have this form is that when data from different conditions are combined the form of the function is masked.

The present experiment then was directed towards obtaining data that would allow us to estimate the values of the parameters in Equation (15) as a function of the significant variables in each experimental condition.

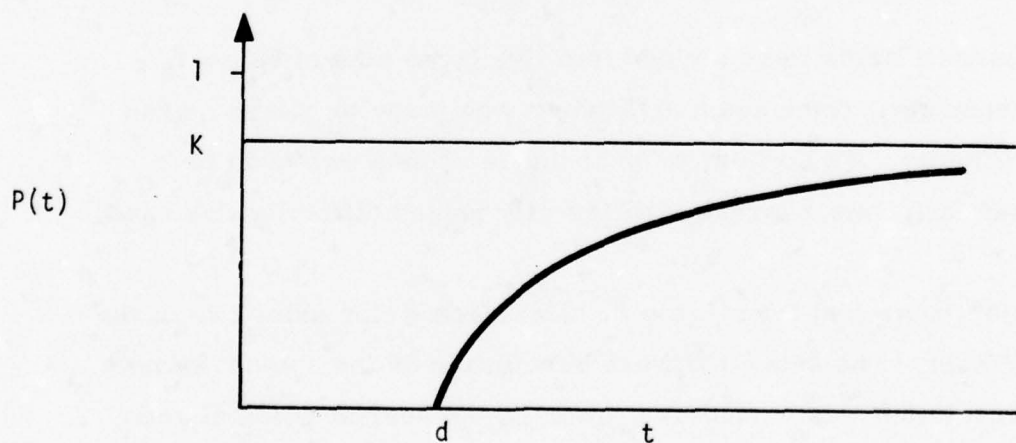


Figure 93. Graph of Equation (15)

SECTION II

METHOD

DESIGN

Each of the 135 subjects searched for targets within 192 search fields or 25,920 trials in all.

The search fields were made in the Honeywell Digital Image Processing Facility from 78 target pictures and 24 background scenes. For each scene a sample of 20 targets was selected. By embedding each of these targets individually in that scene, 20 search fields were created. Thus, there were 480 search fields altogether. The targets were first subjectively graded in difficulty using size and shape as major bases. This was done so that all of the samples of 20 targets for each scene would be roughly equivalent to one another in difficulty.

The 480 search fields were divided into five large sets of 96 search fields in each set. Once again an attempt was made to balance these sets in difficulty. To do this, each of the 24 scenes was used four times, and each time a target of differently judged difficulty was used.

Each subject searched for targets in 192 search fields using two of the five large sets. The sets used were balanced over the subject sample. Display resolution was varied by projecting the search fields slightly

out of focus for half the subjects. The main features of the design are shown in Figure 94.

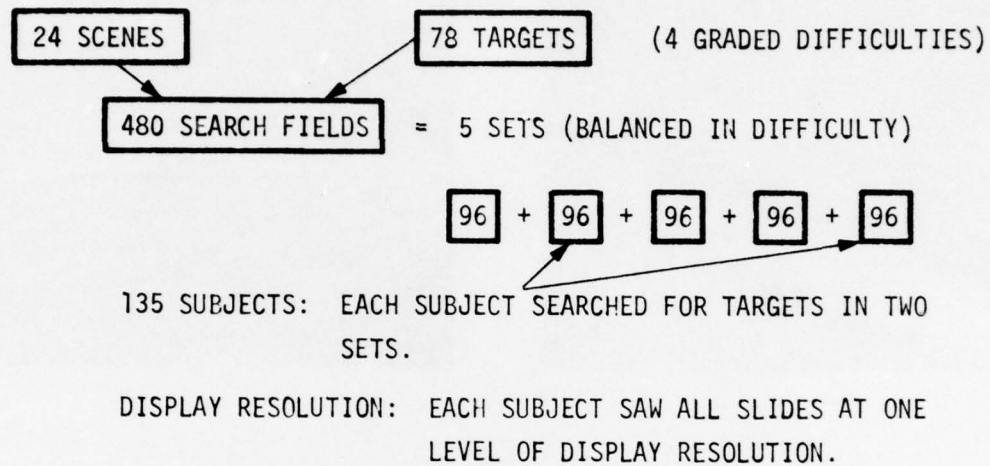


Figure 94. Main Features of the Experimental Design

STIMULI

The 24 scenes (Figure 95) were selected from a file of more than 400 transparencies from an earlier NVL contract. We tried to select as broad a sample as possible from that file. Then, within each scene from one to three "homogeneous" search regions were defined. These regions were selected to have more or less the same texture throughout, and they would be the search regions used in the main experiment.



Figure 95. Four Scenes

There were 78 targets. The number of targets within each category is shown in parentheses.

- Tank (14)
- Truck (20)
- Jeep (14)
- Van (11)
- Gun (9)
- Tractor (7)
- Car (1)
- Small Truck (1)
- Car and Trailer (1)

Of these, 26 were direct photographs of the target, while the other 52 were photographs of the monitor from an infrared sensor in the 3 to 5 μ range. Some target examples are shown in Figure 96.

Each of the 480 search field transparencies was created by embedding by computer a target of appropriate size at the specified location in the scene. Each scene was digitized as an 800 x 800 pixel array and written out on film in a 40 x 40 mm format. The average target size was 8 x 15 cells in the scene. About 25 percent of the targets had a maximum dimension of 10 cells or less. The embedding process is described in greater detail in Appendix B.

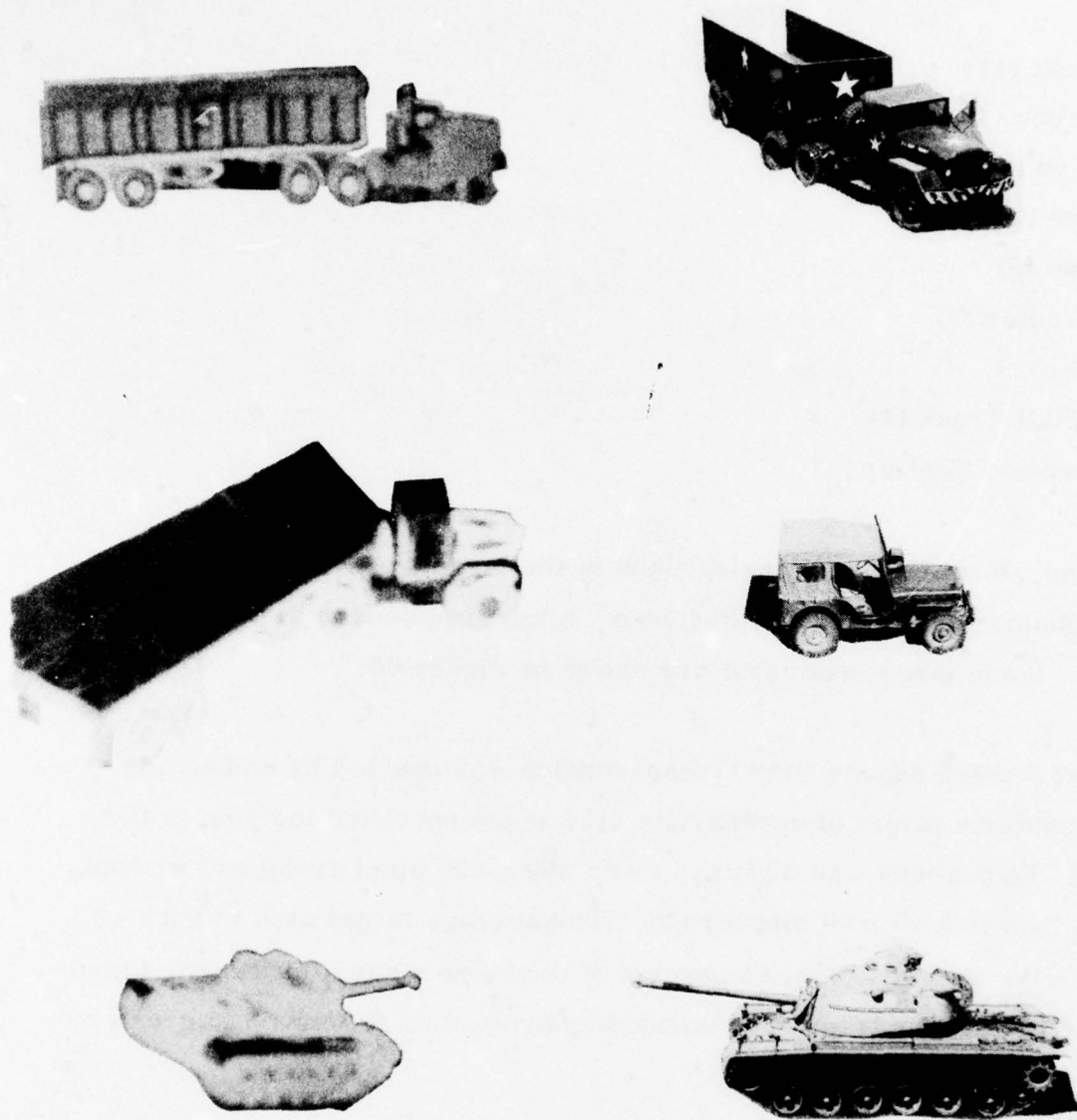


Figure 96. Six Targets

DIFFICULTIES

When the output transparency was inspected it often looked wrong. The target size was wrong or it was badly placed. Therefore, many of the pictures were remade. Probably twice as many pictures as were actually used were made. Many of the final pictures were good, while others were not as realistic. The major flaw was that the luminance range in the target was less than we would have liked. We could, of course, have adjusted the target luminance range, but time and money constraints were already being stretched too much. Therefore, we decided to use the stimulus set we had and not try to improve it further. Another flaw for some of the pictures was that the target was the darkest object in the picture. This sometimes was unrealistic. However, this does resemble the case for the infrared sensor, where the target often is the brightest object in the scene. We did succeed in eliminating any artifacts relating to the embedding process itself. There were no outlines, bright edges, or any other cues which might give the target away.

PROCEDURES

The 135 male and female subjects were students at vocational schools and were paid for their participation. There were from 12 to 14 subjects in each experimental session which lasted two to two and one-half hours. Subjects were trained in the first hour, with the second hour used for the main experimentation.

The slides were presented on the front surface of a screen using a Kodak Carrousel 650 H projector with a 125 mm focal length lens.

The subjects were arranged so that they all could see the screen, but some were more favorably seated than others. The distance ranged from about 1.5 to 3 meters. The projected fields were about 130 x 130 cm square.

The out-of-focus condition was achieved by projecting the standard Air Force 3-bar test pattern and defocusing so that the 4-6 pattern was barely visible with close viewing. When the scene was in focus the 5-4 pattern was visible. The resolution of the projected slides for the two conditions was 45.6 and 28.5 line pairs/mm, as measured on the slide --a ratio of 1.6 to 1. Although we might have defocused the image somewhat more, it would have made an already difficult task too much more so.

TRAINING

During training the subjects were first shown examples of the targets and scenes. They were then allowed to search for a series of targets under conditions closer and closer to that in the main experiment. The experimenter reviewed each search trial and attempted to show where the target was and what signatures allowed it to be identified as such. The search fields used for training a given subject group would be one of the three sets not used for the main trials. The training instructions are given in Appendix C.

MAIN TRIALS

There were 192 main trials. Each trial lasted 15 seconds, with the first 10 seconds allotted for search and at least five seconds for recording data. The subject's task was to find the target as quickly as possible, and to write down the time, the type of target, and the location within the region searched. He or she was told that he need not be certain that a given object was the target--only reasonably sure. Furthermore, if he were unsure as to the type of target, then it was appropriate to guess. The time to find was self-determined by observing the position of the second hand on a modified wall clock face (Figure 97). Target location was defined in terms of general position within the region (e.g., middle, upper-left, etc.).

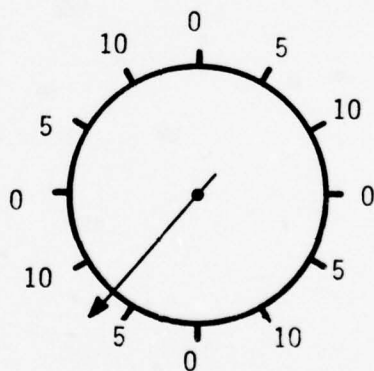


Figure 97. Modified Clock

The experimenter started each trial at the zero position of the second hand on the clock. After each block of 10 trials there was a 15 second break. There would be a longer break when the slide tray was changed.

This procedure required a high level of cooperation from the subjects, and it appears that this requirement was met.

SECTION III

DATA ANALYSIS

OVERVIEW

There were two kinds of data in this study: subject performance data and data relating to the imagery (including the psychometric data). After having obtained the data there were two main steps in its analysis. The first step was curve fitting, where the three parameters in the search equation were estimated for each search condition. Having estimated these parameters, the next step was to use correlation procedures to determine which combination of variables best predicted these parameters and which combination of variables best predicted the probability of finding the target.

ANALYSIS OF THE IMAGERY

There were 480 search fields, each containing one target within an outlined region. For each search field, a wide variety of measures were made to be used in the later analysis. There were basically three types of measures: measures on the target, on the background, and on the relationship between the target and the background. Included in the latter category were some psychometric variables.

Target Measures

Table 15 shows the measures of size and intensity (i.e., luminance) which were used. All the measures except for PER, MXD, and MND were made by computer analysis on the digital representation of the target.

TABLE 15. TARGET MEASURES

Area	TAR
Perimeter	PER
Maximum Dimension	MXD Long dimension of the target
Minimum Dimension	MND Short dimension of the target
$\text{Perimeter}/\sqrt{\text{TAR}}$	
$\text{MXD} \times \text{MND}/\text{TAR}$	
Mean Intensity	MIT
Standard Deviation of Intensity	
Peak Intensity	PIT Maximum inten- sity of any 3 x 3 pixel region on the target
Valley Intensity	VIT Minimum inten- sity of a 3 x 3 pixel region on the target

Background Measures

Table 16 shows the measures of intensity and texture made on each region.

TABLE 16. BACKGROUND MEASURES

Background Region Area	BAR
Mean Intensity	MIB
Standard Deviation Intensity	
Mean Density	
Standard Deviation Density	
Grey-Level Mean	GLM
Grey-Level Contrast	GLC
Angular Second Moment	ASM
Entropy	ENT

These measures were made for five spacings (4, 8, 12, 16, and 20 pixels) coupled with the four major directions: horizontal, vertical, and the two diagonals)

The grey-level measures were defined as follows. First, compute $p_{\delta}(i) = \text{prob} (i = |\phi(x + \Delta x, y + \Delta y) - \phi(x, y)|)$ for the specified spacing, $\delta = (\Delta x, \Delta y)$ where $\phi(x, y)$ is the value (density or intensity) at cell (x, y) .

Then

$$\text{Mean} \quad \text{GLM} = \sum_{i=0}^{N-1} i p_{\delta}(i)$$

$$\text{Contrast} \quad \text{GLC} = \sum_{i=0}^{N-1} i^2 p_{\delta}(i)$$

$$\begin{array}{ll} \text{Angular Second} & \text{ASM} = \sum_{i=0}^{N-1} p_{\delta}^2(i) \\ \text{Moment} & \end{array}$$

$$\begin{array}{ll} \text{Entropy} & \text{ENT} = -\sum_{i=0}^{N-1} p_{\delta}(i) \ln(p_{\delta}(i)) \end{array}$$

The grey-level statistics are typically calculated using the intensities (corresponding to luminances) at each pixel. However, the random variation in intensity is often a function of the overall intensity level. Therefore, brighter scenes would have higher valued grey-level statistics. That is, the values of the grey-level statistics would be highly correlated with the intensity of the background. By using the densities of each pixel, this correlation was eliminated.

Measures on the Search Field--Target in Background

There are a variety of measures of the relationship of the target to the background that should relate to search performance. We have divided them into five classes:

- Target size--Search region size. This should relate to the time required to scan the search region.
- Target size--Display resolution. This relates to the difficulty in recognizing the target.
- Target size--Texture measures. This relates to the embedding of the target in the background texture.

- Target intensity--Scene intensity. This is a measure of the visibility of the target.
- Psychometric--This is judgments of scene features and overall difficulty.

Target Size--Search Region. The ratios of all four measures of target size to background region area (BAR) were used:

TAR/BAR

PER/BAR

MXD/BAR

MXD/\sqrt{BAR}

Target Size--Display Resolution. The ratios of all four measures of target size to display resolution (RES) were used:

\sqrt{TAR}/RES

PER/RES

MXD/RES

MND/RES

Target Size--Texture Measures. The ratios of all four measures of target size to all 20 of the grey-level statistics measures of texture were used. There were 80 such ratios. We found that for most regions the grey-level statistics were independent of direction. Therefore, average overall directions were used.

Target Intensity--Scene Intensity. There were a number of contrast measures used in the analysis. First, there were the three natural

measures of target contrast with the immediate background intensity (IMB) using the mean, peak, and valley target intensities:

$$|MIT - IMB| / \text{MAX}(MIT, IMB) \quad \equiv \quad C_L$$

$$|PIT - IMB| / \text{MAX}(PIT, IMB)$$

$$|VIT - IMB| / \text{MAX}(VIT, IMB)$$

Another similar and often-used measure is $|MIT - IMB| / IMB$. In the present study this would be almost identical to luminance contrast (C_L) since the background was almost always of higher intensity than the target.

There were then the three corresponding measures using the mean intensity of the entire background region. This calculation is made more readily since only one is needed for each background.

$$|MIT - MIB| / \text{MAX}(MIT, MIB)$$

$$|PIT - MIB| / \text{MAX}(PIT, MIB)$$

$$|VIT - MIB| / \text{MAX}(VIT, MIB)$$

Other contrast measures were also tried. For example, the six simple intensity ratios were used.

$$MIT/IMB \quad MIT/MIB$$

$$PIT/IMB \quad PIT/MIB$$

$$VIT/IMB \quad VIT/MIB$$

The first of the above is a commonly used measure of contrast. We use the symbol, C_L , luminance contrast.

Psychometrics

Five psychometric measures were used. First, there were two measures of judged difficulty:

Judged difficulty of locating the target (JDL)

Judged difficulty of identifying the target (JDI)

These judgments were made for all 480 search fields by the two experimenters who were highly familiar with all the search fields. JDL was the judged difficulty of finding the target object irrespective of the problems in identifying it. JDI was the judged difficulty of identifying the target when looking directly at it. A five-point scale was used for JDL and a three-point scale for JDI.

Three ratings of texture were used. A group of seven students independently rated all 46 search regions from the 24 scenes with respect to the following features. A five-point scale was used for each feature.

- Clutter--How many other objects are there in the region?
- Number of Confusing Objects--How many target-like objects are there in the region?
- Homogeneity--How much does one part of the region look like any other part?

SUBJECT PERFORMANCE DATA

Each of the 135 subjects responded to 192 search fields. There were a number of possible outcomes for each trial:

- The target was found and correctly identified in from 1 to 10 seconds;
- The target was found, but incorrectly identified in from 1 to 10 seconds; or
- The target was not found.

The difficulty in scoring a trial is the scorer's uncertainty as to whether or not the subject has found that object we call the target. In fact, there often were many other objects that well could have been targets but were not. Of course, the subject's recorded location was a major basis for deciding whether the target was found. If the location was at or close to the true location, then the trial was scored as a correct one unless the identification was grossly different from the true target (e.g., a tank identified as a gun). If the identification was correct, then somewhat more leeway in location was permitted.

It is obvious that considerable error was possible with this procedure. In the future, we might overcome this difficulty by using a fine coordinate grid projected over the scene. However, the small size of some of the search regions can create additional difficulties.

For each of the 960 search conditions (480 search fields x 2 display resolutions) two search time distributions were derived: the cumulative probability of identifying the target and the cumulative probability of detecting (or locating) the target. Each search time distribution was based on from 24 to 28 subject responses.

PARAMETER ESTIMATION--CURVE FITTING

A combination of grid-search and Newton's method was used to fit the best curve of the form:

$$p(t) = \begin{cases} K (1 - e^{-\lambda(t-d)}) & \text{for } t \geq d \\ 0 & \text{otherwise} \end{cases}$$

where $0 < K \leq 1$, $\lambda > 0$ and $d > 0$. d was estimated to be the highest (integer) time value for which $p(t) = 0$ in the data. A grid-search (with .02 fineness) was made in the interval

$$K, \lambda \in (0, 1)$$

for the minimum (on the grid) of

$$\epsilon^2 = \sum_{k=1}^{10} (p_k - p(k))^2$$

where p_k is a data point. If, as often occurred, the minimum was on the upper boundary for K , K was set to 1.0 and Newton's method was used to further minimize ϵ^2 with respect to λ . Otherwise, Newton's method for functions of more than one variable was used to further reduce the grid-search estimate with respect to both K and λ .

A total of 157 measurements or combinations of measurements were used for each of the 1,920 data events. These included criterion measures (K , λ , etc.), scene measures, target measures, and psychometric variables. Pearson product-moment correlation coefficients were calculated for each pair of measures. This matrix was then examined for various relationships:

1. Which variables or subsets of variables appeared to be most strongly related to each criterion variable?
2. How did various measurements relate to one another?
3. Which "objective" measures were related to the psychometric variables?

RESULTS

Our objective has been to predict to five main criteria: the three parameters in the model expressed by Equation (15) and the two search times distributions, the probability of locating the target and the probability of identifying it. The first outcome of the regression analysis was that we could predict points of the search time distributions better than we could any of the three parameters. The second main finding was that the same or similar predictor variables were found to be most useful for all the criteria.

An Initial Comparison of Target Location and Identification

We found that the probability of identifying the target is highly correlated with the probability of locating it, with the correlation over the 960 search conditions equal to 0.72. The correlations between the two main performance measures and each predictor variable typically did not differ greatly. This should be expected since they are both based on the same responses. In some cases, there were differences but those differences were small. Most notable differences were for the maximum target dimension (MXD) and MXD/RESOLUTION, which were more strongly related to target identification than to detection. Judged clutter, number of confusing objects, and homogeneity were more strongly related to detection than to identification.

Predicting Target Location

Predicting P(10) and P(1)-- The best objective prediction of P(10) was made from two parameters, luminance contrast (C_L) and texture contrast (C_T) where

$$C_T = \text{MXD} / \sqrt{1 + \text{GLM}_4^2}$$

GLM_4 is the grey-level mean calculated over a spacing of four pixels. We use this form of the denominator because when we simply divide by GLM_4 , with very low values of GLM_4 the correlation coefficients and the resulting prediction equations become greatly distorted.

The following equation is the best linear combination of the two parameters. The correlation between \hat{P} and $P(10)$ is 0.54.

$$\hat{P} = .44C_L + .043 C_T + .19$$

Moreover, virtually no improvement results by adding other variables to the prediction equation. For the range of values studied here we found that display resolution improved the prediction slightly, but search field area did not help at all. Resolution and area did correlate with $P(10)$ but only moderately so. The relationship between $P(10)$ and C_L and C_T is shown in Figure 98.

If contrast and target size were used as predictors, then the correlation between $P(10)$ and the linear combination of size and contrast became somewhat lower, i.e., 0.47.

The best prediction of $P(1)$ (early detection performance) is made from C_L , C_T and resolution. Now we find that resolution adds greatly to the predictive power of C_L and C_T . The correlation of the prediction equation with $P(1)$ was 0.69. Thus, $P(1)$ was better predicted than $P(10)$.

Judged difficulty of target location (JDL) was an even better predictor of $P(10)$ than the "objective" predictors. The correlation between JDL and $P(10)$ for target location was 0.76. This correlation also could not be improved by adding judged difficulty of identification (JDI) or any other variables. The relationship between $P(10)$ and JDL is shown in Figure 99.

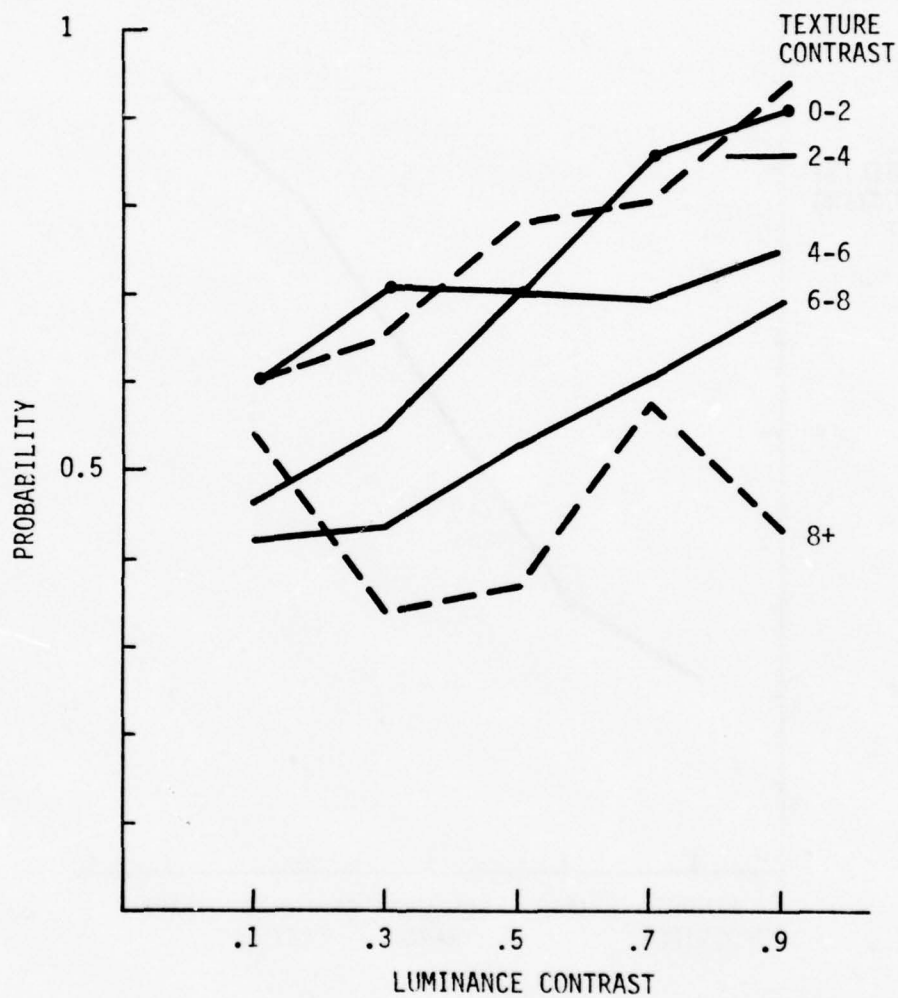


Figure 98. Probability of Locating Target as a Function of Luminance Contrast and Texture Contrast

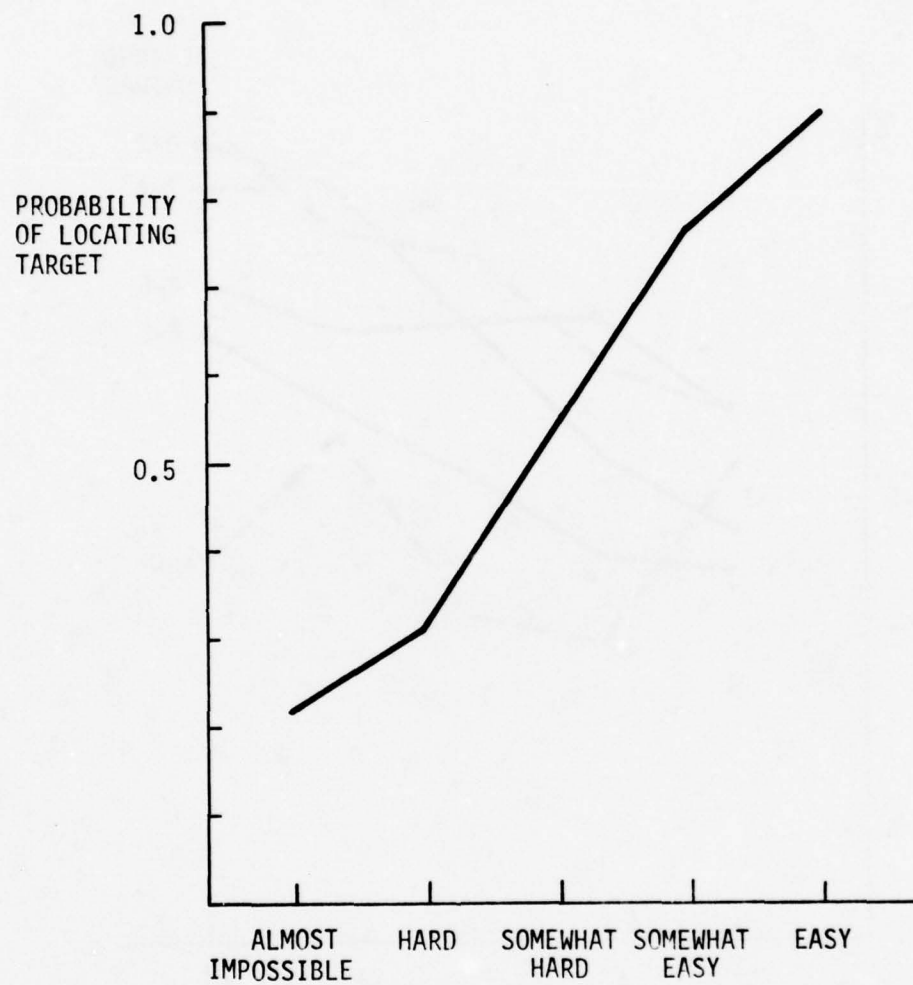


Figure 99. Judged Difficulty of Locating Target

JDL was also the best predictor of P(1). However, adding resolution to the prediction equation aided greatly, for the prediction equation correlated 0.75 with P(1). We see that P(1) is predicted about as well as P(10). We found that the judgment of clutter and homogeneity had little predictive utility, as compared with JDL and JDI.

Predicting Target Identification

Predicting P(10) and P(1)--The best objective prediction of P(10) also was made from C_L and C_T . The prediction equation is similar to that for the detection distribution:

$$\hat{P} = .44 C_L + .041 C_T - .06$$

Adding resolution improved the correlation from 0.56 to 0.59. Similarly, to predict P(1) it is best to use C_T , C_L , and resolution. The correlation of the prediction equation with P(1) is then 0.72. However, we find that we can do almost as well by using MXD instead of C_T . In other words, to predict the recognition time distribution, it is sufficient to ignore texture and use target size, target-background contrast, and resolution. We see also that P(1) here is better predicted than P(10).

To predict to P(10) from the subjective measures we found that JDI was better than JDL. In contrast to predicting the target location distribution, we found that JDI and JDL together were an improvement over JDI alone. Adding resolution improved the prediction just slightly, with the correlation now becoming 0.74.

Similarly, to predict $P(1)$ it was best to use JDI, JDL and resolution. The prediction then correlates 0.75 with $P(1)$.

Predicting the Parameters in Equation (15)--The most important result is that at this time we cannot predict to the three parameters very well. We find that the variables that predict the best for $P(10)$ and $P(1)$ are also best for the three parameters.

The results are summarized in Table 17 which shows the correlation between combinations of specified sets of predictor variables and the three parameters in Equation (15) as well as $P(1)$ and $P(10)$.

TABLE 17. CORRELATION BETWEEN A LINEAR COMBINATION OF PREDICTOR VARIABLES AND PERFORMANCE CRITERIA

Performance Distribution	Predictors	Criteria				
		K	λ	d	$P(1)$	$P(10)$
Target Location	$C_L + C_T$.29	.11	.38	.51	.54
	$C_L + C_T + \text{RES}$.29	.10	.43	.69	.57
	JDL	.42	.19	.57	.57	.76
	JDL + RES	.42	.19	.60	.75	.78
Target Identification	$C_L + C_T$.42	.06	.42	.63	.56
	$C_L + C_T + \text{RES}$.43	.06	.46	.72	.59
	JDI + JDL	.51	.11	.61	.70	.72
	JDI + JDL + RES	.52	.10	.64	.75	.74

We see that the correlations with the three parameters are low; we need to determine how to improve these predictions. Especially needed is improvement in predicting λ , which corresponds to the rise time in the search time distribution, where the correlation is extremely low.

COMMENTS

It is important to make two points here. First, we have been using the correlation of a linear combination of predictor variables with $P(10)$, or any of the other criteria, as a way to compare different sets of predictors. However, it would be a mistake to use the size of the correlation coefficient as a way of judging the quality of the prediction system. The reason is that we can always obtain high correlations by using broad ranges of variables. In the present study, by using a wider range of display resolution and a wider range of target sizes, we could easily have obtained correlations of 0.90 and higher. But what we were attempting to do was determine which variables best predicted performance. This we were able to do by isolating the effects of non-homogeneity of the overall search field.

The second point relates to the many measures on the target and background and how they were reduced to a very few. It happens that the several variables of a given kind are correlated among themselves. The various size measures were highly correlated; the several contrast measures were highly correlated; and, the many texture measures were highly correlated. Therefore, for each class of variables we selected that which best predicted the criteria. Thus, MXD was the best measure of size and GLM_4 was the best measure of texture. However, GLM_8 was almost

as good as GLM_4 for measuring texture and there might be reasons for using it instead. For example, GLM_8 might be easier to measure under some circumstances. At the beginning of the study we had considered making other measures of texture, but we decided against this since other researchers found that using these measures provided no benefit. In any event, the other measures would be highly correlated with those we had been using.

REFERENCES

1. "Automated Image Enhancement Techniques for Second Generation FLIR," Contract No. DAAG53-76-C-0195, First Interim Report, Honeywell Systems and Research Center, Defense Documentation Center No. AD-AO-35627, December 20, 1976.
2. "Automated Image Enhancement Techniques for Second Generation FLIR," Contract No. DAAG53-76-C-0195, Second Interim Report, Honeywell Systems and Research Center, Defense Documentation Center No. AD-AO-41034, May 15, 1977.
3. "Automated Image Enhancement Techniques for Second Generation FLIR," Contract No. DAAG53-76-C-0195, Interim Report, Honeywell Radiation Center, Defense Documentation Center No. AD-AO-43572, August 1977.
4. Morgan, D. R., "Analog Sorting Network Ranks Input by Amplitude and Allows Selection," Electronics Design, pp. 72-73, January 1973.
5. Moving Image Processor, Honeywell Radiation Center, October 1976.
6. Digital Operations for Image Registration, Honeywell Radiation Center, October 1976.
7. "Night Vision Laboratory Static Performance Model for Thermal Viewing Systems," U.S. Army Electronics Command, ECOM-7043, April 1973.
8. Lloyd, J. M., Thermal Imaging Systems, New York: Plenum Press, 1975.
9. Mood, A. M., "On the Asymptotic Efficiency of Certain Non-Parametric Two-Sample Tests," Annals of Mathematical Statistics, 1954, 25, pp. 514-522.
10. Monk, T. H., "Sequential Effects in Visual Search," Acta Psychologica, 38, 1974, pp. 315-321.

REFERENCES (concluded)

11. Monk, T.H. and Brown, B., "The Effect of Target Surround Density on Visual Search Performance," Human Factors, 1975, 17, pp. 356-360.
12. Bloomfield, J. R., Graf, C. P., and Graffunder, K., "Visual Factors in Target Disruption," Honeywell Inc., Minneapolis, MN, Final Technical Report under contract No. DAAK02-75-C-0055, 1975.
13. Scanlan, L.A., Hershberger, M. L., and Herman, J.A., "Contrast Enhancement Using Local Area Brightness and Gain Control," Proceedings of the Human Factors Society 21st Annual Meeting, (Ed. A.S. Neal and R. F. Palasek), San Francisco, CA, 1977, pp. 231-234.
14. Rhodes, Fen, "Predicting the Difficulty of Locating Targets from Judgments of Image Characteristics," AMRL-TDR-64-19, 1964.
15. Weska, J.S., Dyer, C. R., and Rosenfeld, A., "Comparative Study of Texture Measures for Texture Classification," ITLL Trans. on Systems, Man, and Cybernetics, April 1976.

APPENDIX A

PERFORMANCE PREDICTION PROGRAM OUTPUT

INPUT DATA

RUN NUMBER 1

YOUR SPECTRAL BAND IS 8.000 TO 12.000 MICRONS

OPTICS

DIAMETER 6.001 INCHES
 F-NUMBER 3.333
 FOCAL LENGTH 20.000 INCHES
 AVG. OPTICAL TRANSMISSION .550
 WAVELENGTH FOR DIFFRACTION 10.000 MICRONS
 GEOMETRIC BLUR SPOT SIZE -0.000 MRAD.

DETECTOR

HORIZONTAL IFOV .050 MRAD.
 VERTICAL IFOV .050 MRAD.
 DETECTORS IN PARALLEL 630.
 DETECTORS IN SERIES 2.
 DETECTOR SIZE .00100 INCHES
 PEAK D* 1.00 (1E10)CM-SORT(HZ)/WATT
 MEASURING FREQUENCY OF D* 5000. HERTZ
 COLD SHIELD ANGLE 8.627 DEGREES
 LIMITING NOISE DETECTOR
 DETECTOR RESPONSE, 3-DB POINT .100E+06 HERTZ

SCANNER

FRAME RATE 60.000 FRAMES/SECOND
 SCAN EFFICIENCY .935
 OVERSCAN RATIO 1.000

ELECTRONICS

PREAMP. LOW FREQ 3-DB CUT-ON .200 HERTZ
 AMPLIFIER, 3-DB POINT 0.000 HERTZ
 E/O LED WIDTH 0.00000 MRAD.
 E/O LED LENGTH 0.00000 MRAD.
 APERTURE CORRECTION AMPLITUDE 0.000 HERTZ
 APERTURE CORRECTION FREQUENCY 0.

INPUT DATA

DISPLAY

TYPE	CRT DISPLAY
X SPOT SIZE	.004 MRAD.
Y SPOT SIZE	.004 MRAD.
AVERAGE BRIGHTNESS	10.000 FT. LAMBERTS

SYSTEM

HORIZONTAL FOV	1.800 DEGREES
VERTICAL FOV	1.400 DEGREES
MAGNIFICATION	22.222
WFOV/NFOV	1.000
NOISE EQUIV. DELTA T	-0.000 DEGREES C

STABILIZATION

SYSTEM STATE	STABILIZED
X VIBRATION CONSTANT	0.00
Y VIBRATION CONSTANT	0.00

STANDARD INPUTS

EYE INTEGRATION TIME	.200 SECONDS
THRESHOLD SIGNAL/NOISE	2.250

INPUT DATA

ATMOSPHERIC PARAMETERS

CONDITION	CLEAR	KILOMETERS
VISIBILITY RANGE	23.000	PERCENT
RELATIVE HUMIDITY	46.000	DEGREES C
AIR TEMPERATURE	15.000	

TARGET & BACKGROUND

TARGET LENGTH	5.250	METERS
TARGET WIDTH	2.700	METERS
TARGET DELTA T	1.000	DEGREES C
BACKGROUND TEMPERATURE	12.000	DEGREES C

RANGE REQUIREMENTS

MIN. REQUIRED RANGE FOR RECOG.	3000	METERS
MAX. REQUIRED RANGE FOR RECOG.	12000	METERS
RANGE INCREMENTS FOR RECOG.	1000	METERS
MIN. REQUIRED RANGE FOR DETEC.	5000	METERS
MAX. REQUIRED RANGE FOR DETEC.	20000	METERS
RANGE INCREMENTS FOR DETEC.	1666	METERS

INPUT DATA

TEMPORAL MTF'S

FREQ. (LOG HERTZ)	0.00	3.44	3.80	3.97	4.10	4.19	4.27	4.34	4.40	4.50
DETECT. ROLLOFF MTF (X)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DETECT. ROLLOFF MTF (Y)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ELECTRONIC MTF	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
BOOST MTF	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

SPATIAL COMPONENTS MTF'S

FREQ. (CYC./MRAD.)	0.00	1.50	3.00	4.50	6.00	7.50	9.00	10.50	12.00	13.50
OPTICS MTF (X)	1.00	.34	.76	.65	.54	.44	.34	.26	.17	.10
OPTICS MTF (Y)	1.00	.44	.76	.65	.54	.44	.34	.26	.17	.10
VIDICON MTF (X)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
VIDICON MTF (Y)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
STABILIZATION MTF (X)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
STABILIZATION MTF (Y)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

NOISE POWER SPECTRUM (VOLTS/SQRT(HZ))

FREQ. (LOG HERTZ)	1.00	1.30	2.00	2.70	3.00	4.00	5.00	6.00	0.00	0.00
POWER (TIMES 1E-9)	64.00	32.70	16.00	7.12	4.00	1.00	1.00	1.00	0.00	0.00

D* OF DETECTOR (CM.*SQRT(HZ.)) /WATT

WAVELENGTH (MICRONS)	1.00	1.50	2.00	2.50	3.00	3.50	4.00	4.50	5.00	5.50
D* (TIMES 1E10)	.23	.34	.50	.58	.73	.78	.84	1.00	1.00	.25

INPUT DATA

DETECTION + RECOGNITION PROBABILITY DENSITY

DETECTION FREQ.	12.50	5.00	3.00	2.00	1.50	1.00	.75	.50	.25	0.00
BEST RECOG. FREQ.	37.50	15.00	9.00	5.00	4.50	3.00	2.25	1.50	.75	0.00
WORST RECOG. FREQ.	50.00	20.00	12.00	8.00	6.00	4.00	3.00	2.00	1.00	0.00
PROBABILITY	1.00	1.00	1.00	.95	.80	.50	.30	.10	.02	0.00

CCD CIRCUIT 1.0 FOCAL PLANE TEMP= 77.0 K

```

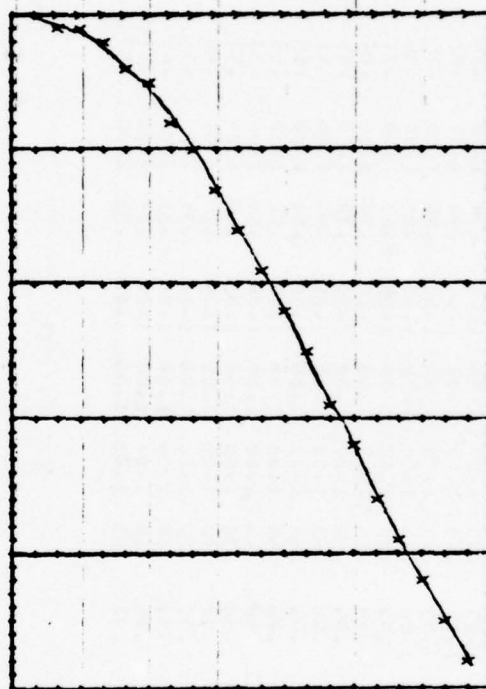
***** INPUT PARAMETERS *****
SN .200E+01. PN .530E+03. TZ .320E-04. TS .508E-07. T .770E+02. WAV2 .120E+02. PH18 -0.
CKT .100E+01. AG .250E-05. CW .200E-12. CN .100E-09. FA .197E+09
TYPE -0. AD .625E-05. MA+D .400E+02. CD .100E-06. QEFF .700E+00. AEX -0. ; GTM -0.
E,IF .519E-10. AG2 -0. ZL -0. ; RBH -0. ; MFE -0. ; EFT -0. ; CL -0.
DC CURRENT (PH18) = .958E-08 AMPS
CCD INPUT CURRENT (NA)=10.561
SCAN VELOCITY IN M/SEC= .15722E+04
PV DIODE REVERSE CURRENT(NA) 1.038
COMPUTED PV PEAK D= 27.69 (IE10)CM-SQRT(HZ)/WATT

```

D= OF DETECTOR VS WAVELENGTH

WAVELENGTH (MICRONS)	8.00	9.44	9.99	9.33	9.78	10.22	10.67	11.11	11.56	12.00
RELATIVE D= CURVE	.57	.70	.74	.78	.81	.85	.89	.93	.96	1.00

FREQ	MTF-X	MTF-Y
0.00	1.000	1.000
1.00	.996	1.000
2.00	.983	1.000
3.00	.963	1.000
4.00	.934	1.000
5.00	.899	1.000
6.00	.856	1.000
7.00	.807	1.000
8.00	.753	1.000
9.00	.694	1.000
10.00	.632	1.000
11.00	.566	1.000
12.00	.498	1.000
13.00	.430	1.000
14.00	.361	1.000
15.00	.293	1.000
16.00	.227	1.000
17.00	.163	1.000
18.00	.102	1.000
19.00	.046	1.000



NOISE POWER SPECTRUM (NORMALIZED)

FREQ/LOG HERTZ	3.59	3.90	4.07	4.20	4.29	4.37	4.44	4.50
NVL NOISE	1.03	.93	.76	.54	.32	.14	.03	.00

DETECTOR-CCD DPEAK (IE10)= 29.663

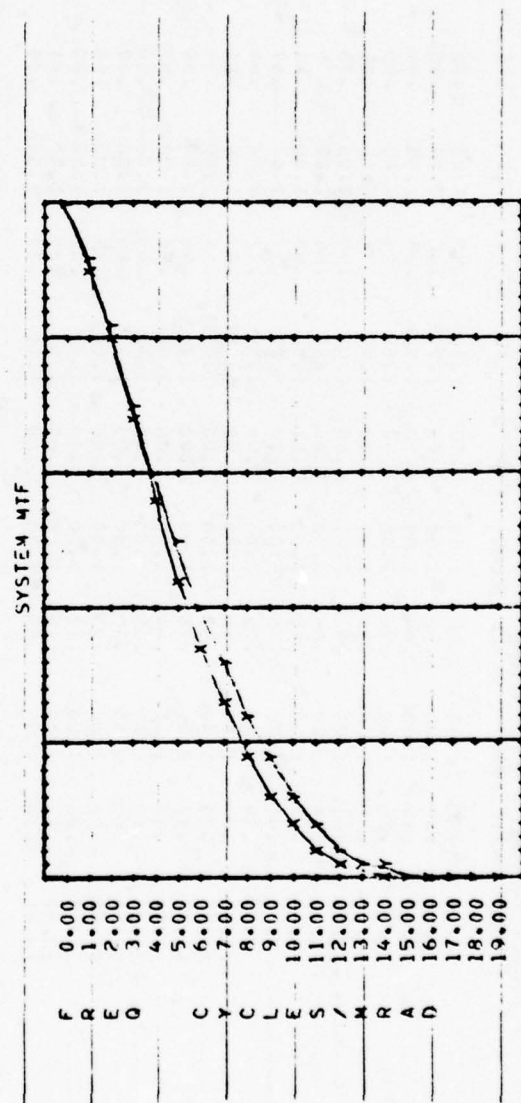
X MODULATION TRANSFER FUNCTIONS

FREQ	OPTIC	GALUR	DETEC	MS&NS	ELECT	BOOST	VIDCN	LED	DS-PLY	LDS	EYE
0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	.92	1.00	1.00	1.00	.99	1.00	1.00	1.00	1.00	1.00	.95
2.00	.84	1.00	.94	1.00	.98	1.00	1.00	1.00	.94	1.00	.91
3.00	.76	1.00	.96	1.00	.96	1.00	1.00	1.00	.96	1.00	.86
4.00	.69	1.00	.94	1.00	.93	1.00	1.00	1.00	.94	1.00	.82
5.00	.61	1.00	.90	1.00	.90	1.00	1.00	1.00	.90	1.00	.78
6.00	.54	1.00	.86	1.00	.86	1.00	1.00	1.00	.86	1.00	.74
7.00	.47	1.00	.81	.99	.80	1.00	1.00	1.00	.82	1.00	.71
8.00	.41	1.00	.76	.99	.75	1.00	1.00	1.00	.77	1.00	.67
9.00	.34	1.00	.70	.99	.69	1.00	1.00	1.00	.72	1.00	.64
10.00	.29	1.00	.64	.99	.63	1.00	1.00	1.00	.67	1.00	.61
11.00	.23	1.00	.57	.99	.56	1.00	1.00	1.00	.61	1.00	.58
12.00	.17	1.00	.50	.98	.50	1.00	1.00	1.00	.56	1.00	.55
13.00	.12	1.00	.44	.98	.43	1.00	1.00	1.00	.50	1.00	.53
14.00	.08	1.00	.37	.98	.35	1.00	1.00	1.00	.45	1.00	.50
15.00	.03	1.00	.30	.97	.29	1.00	1.00	1.00	.40	1.00	.48
16.00	.02	1.00	.23	.97	.23	1.00	1.00	1.00	.35	1.00	.45
17.00	.06	1.00	.17	.97	.17	1.00	1.00	1.00	.31	1.00	.43
18.00	.11	1.00	.11	.96	.11	1.00	1.00	1.00	.27	1.00	.41
19.00	.16	1.00	.05	.96	.05	1.00	1.00	1.00	.23	1.00	.39

Y MODULATION TRANSFER FUNCTIONS
 FREQ OPTIC GORLJR DETEC

FREQ	OPTIC	GORLJR	DETEC	ELECT	VIDCN	LED	USPLY	LOS	EYE
0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	.92	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.95
2.00	.44	1.00	.98	1.00	1.00	1.00	.98	1.00	.91
3.00	.76	1.00	.95	1.00	1.00	1.00	.96	1.00	.86
4.00	.69	1.00	.94	1.00	1.00	1.00	.94	1.00	.82
5.00	.61	1.00	.90	1.00	1.00	1.00	.90	1.00	.78
6.00	.54	1.00	.85	1.00	1.00	1.00	.86	1.00	.74
7.00	.47	1.00	.81	1.00	1.00	1.00	.82	1.00	.71
8.00	.41	1.00	.75	1.00	1.00	1.00	.77	1.00	.67
9.00	.34	1.00	.70	1.00	1.00	1.00	.72	1.00	.64
10.00	.29	1.00	.64	1.00	1.00	1.00	.67	1.00	.61
11.00	.23	1.00	.57	1.00	1.00	1.00	.61	1.00	.58
12.00	.17	1.00	.50	1.00	1.00	1.00	.56	1.00	.55
13.00	.12	1.00	.44	1.00	1.00	1.00	.50	1.00	.53
14.00	.08	1.00	.37	1.00	1.00	1.00	.45	1.00	.50
15.00	.03	1.00	.30	1.00	1.00	1.00	.40	1.00	.48
16.00	-.02	1.00	.23	1.00	1.00	1.00	.35	1.00	.45
17.00	-.06	1.00	.17	1.00	1.00	1.00	.31	1.00	.43
18.00	-.11	1.00	.11	1.00	1.00	1.00	.27	1.00	.41
19.00	-.16	1.00	.05	1.00	1.00	1.00	.23	1.00	.39

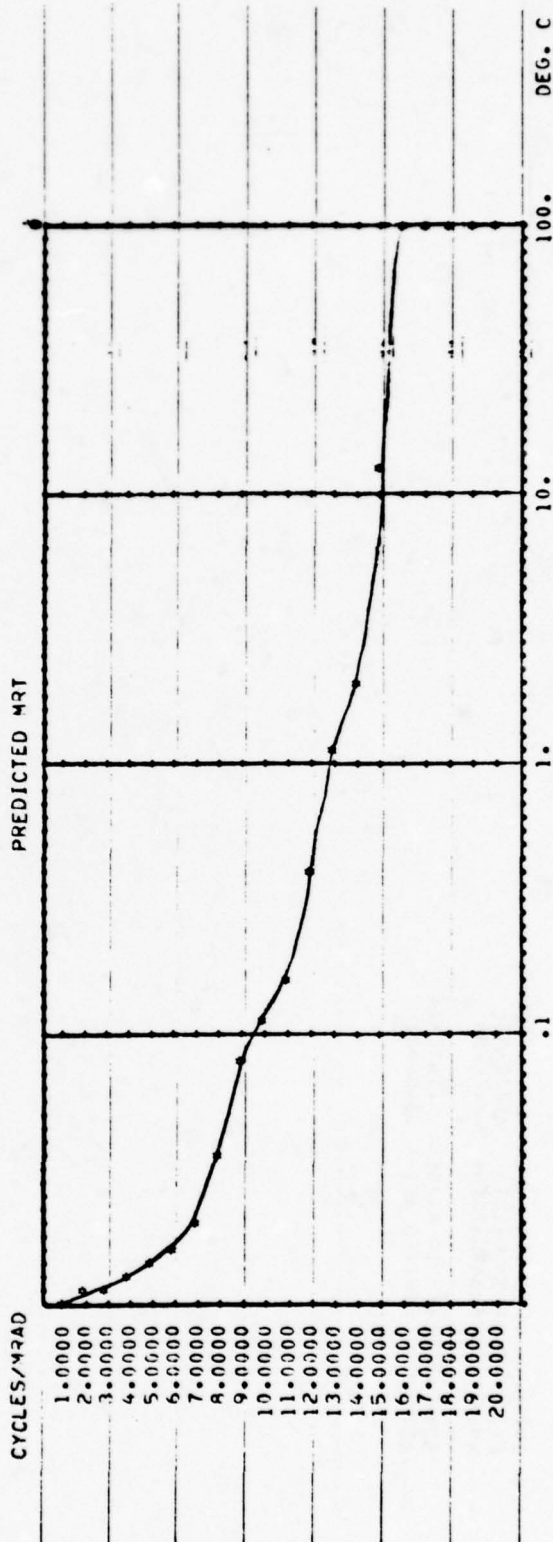
PREDICTED SYSTEM MTF		PREDICTED NOISE FILTERING MTF	
FREQ	X MTF	FREQ	Y MTF
0.00	1.00	0.00	1.00
1.00	.90	1.00	.97
2.00	.80	3.00	.89
3.00	.68	5.00	.76
4.00	.56	7.00	.61
5.00	.45	9.00	.46
6.00	.34	11.00	.32
7.00	.25	13.00	.20
8.00	.18	15.00	.11
9.00	.12	17.00	.05
10.00	.08	19.00	.01
11.00	.04	20.00	.01
12.00	.02	22.00	.02
13.00	.01	24.00	.02
14.00	.00	26.00	.02
15.00	.00	28.00	.04
16.00	.00	30.00	.02
17.00	.00	32.00	.01
18.00	.00	34.00	.01
19.00	.00	36.00	.01



INTEGRAL OF D-STAR*W-PRIME= .250E+08
EXACT NOISE BANDWIDTH= .127E+05
WHITE NOISE BANDWIDTH= .247E+05
SCAN VELOCITY IN M/SEC= .157E+04
DET NOISE LIMITED NET= .340E-01

PREDICTED MINIMUM RESOLVABLE TEMPERATURE

FREQ	X MRT	Y MRT	XL MRT	YL MRT	45MRT
.100E+01	.130E-02	.129E-02	.411E-02	.409E-02	.579E-02
.200E+01	.293E-02	.289E-02	.554E-02	.546E-02	.920E-02
.300E+01	.517E-02	.499E-02	.744E-02	.711E-02	.131E-01
.400E+01	.829E-02	.791E-02	.131E-01	.124E-01	.180E-01
.500E+01	.131E-01	.118E-01	.185E-01	.167E-01	.249E-01
.600E+01	.204E-01	.177E-01	.264E-01	.229E-01	.349E-01
.700E+01	.324E-01	.265E-01	.397E-01	.317E-01	.500E-01
.800E+01	.527E-01	.404E-01	.589E-01	.451E-01	.742E-01
.900E+01	.901E-01	.635E-01	.950E-01	.669E-01	.116E+00
.100E+02	.156E+00	.100E+00	.156E+00	.100E+00	.185E+00
.110E+02	.294E+00	.168E+00	.280E+00	.160E+00	.323E+00
.120E+02	.621E+00	.312E+00	.567E+00	.285E+00	.634E+00
.130E+02	.140E+01	.503E+00	.123E+01	.529E+00	.134E+01
.140E+02	.390E+01	.140E+01	.329E+01	.118E+01	.350E+01
.150E+02	.145E+02	.536E+01	.151E+02	.438E+01	.157E+02
.160E+02	.160E+08	.160E+08	.126E+08	.126E+08	.179E+08
.170E+02	.170E+08	.170E+08	.130E+08	.130E+08	.184E+08
.180E+02	.180E+08	.180E+08	.134E+08	.134E+08	.190E+08
.190E+02	.190E+08	.190E+08	.134E+08	.138E+08	.195E+08
.200E+02	.155E+03	.200E+08	.110E+03	.141E+08	.141E+08

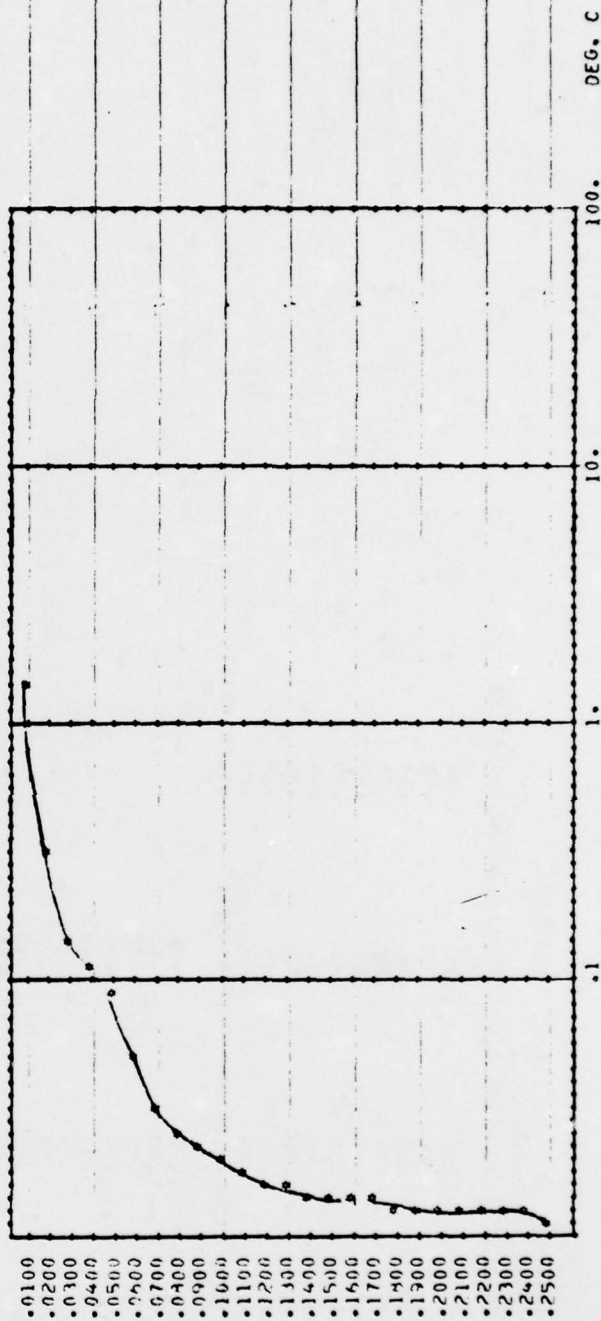


PREDICTED MINIMUM DETECTABLE TEMPERATURE
(TARGET SIZE)°-1

.100E+03	.220E+01
.500E+02	.556E+00
.333E+02	.251E+00
.250E+02	.144E+00
.200E+02	.945E-01
.167E+02	.676E-01
.143E+02	.513E-01
.125E+02	.407E-01
.111E+02	.333E-01
.100E+02	.290E-01
.909E+01	.241E-01
.833E+01	.210E-01
.769E+01	.186E-01
.714E+01	.166E-01
.667E+01	.150E-01
.625E+01	.137E-01
.589E+01	.126E-01
.556E+01	.116E-01
.526E+01	.108E-01
.500E+01	.100E-01
.476E+01	.940E-02
.455E+01	.884E-02
.435E+01	.833E-02
.417E+01	.788E-02
.400E+01	.747E-02

PREDICTED MDT

TARGET SIZE (MYAD)



RECOGNITION PERFORMANCE
 TARGET DELTA TEMPERATURE IS 1.00 DEGREES C
 TEMPERATURE DEPENDANT PERFORMANCE
 RANGE ATM TRANS XCONSERVATIVE YCONSERVATIVE 45CONSERVATIVE

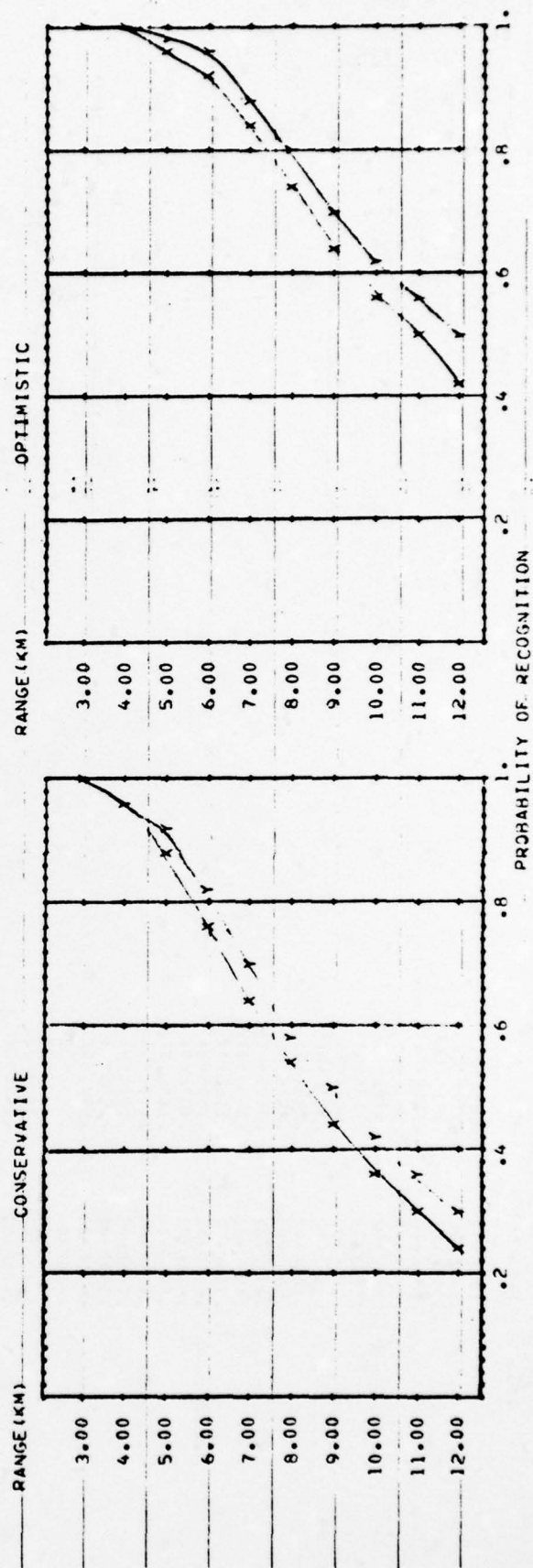
3.00	.80	1.00	1.00	1.00
4.00	.76	.96	.97	.95
5.00	.73	.88	.92	.87
6.00	.71	.77	.82	.76
7.00	.68	.63	.69	.62
8.00	.66	.53	.59	.53
9.00	.64	.44	.51	.44
10.00	.62	.36	.42	.36
11.00	.60	.30	.35	.29
12.00	.58	.25	.29	.24

45OPTIMISTIC

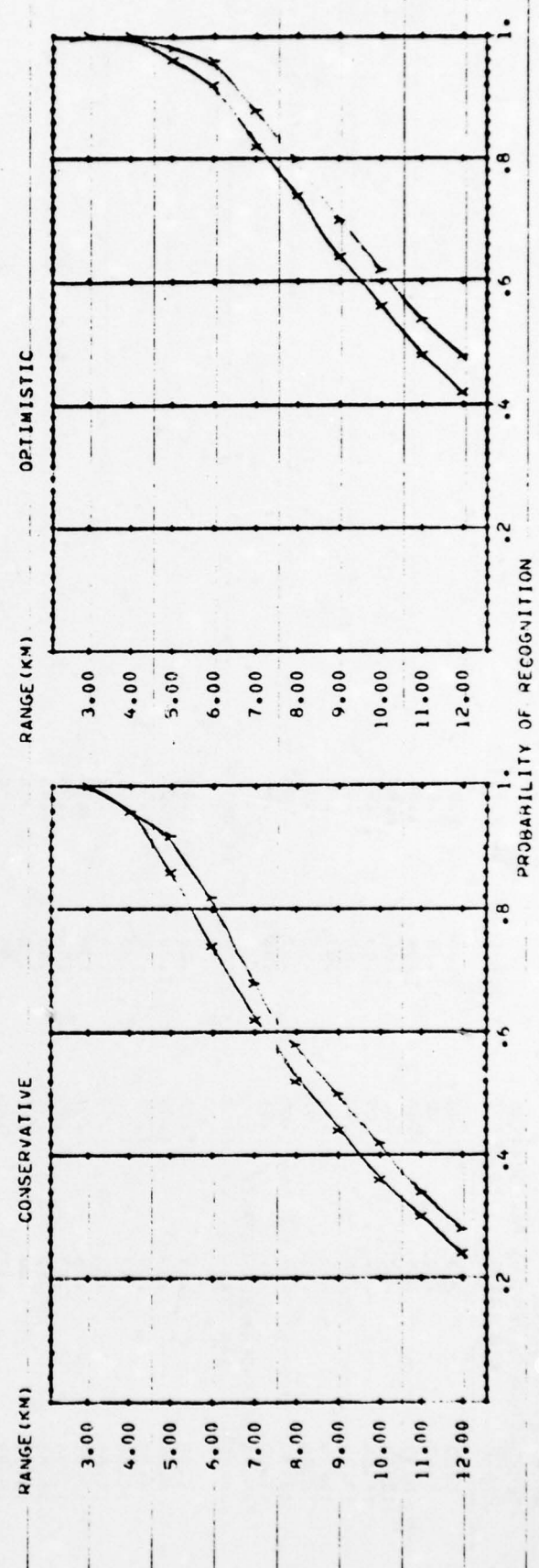
3.00	.80	1.00	1.00	1.00
4.00	.76	1.00	1.00	1.00
5.00	.73	.97	.98	.97
6.00	.71	.93	.95	.92
7.00	.68	.84	.88	.83
8.00	.66	.74	.81	.74
9.00	.64	.64	.71	.64
10.00	.62	.56	.62	.56
11.00	.60	.50	.55	.49
12.00	.58	.43	.49	.42

CONSERVATIVE

OPTIMISTIC



RECOGNITION PERFORMANCE		1.00 DEGREES C	
TARGET DELTA TEMPERATURE IS			
POWER DEPENDANT PERFORMANCE			
RANGE	ATM TRANS	XCONSERVATIVE	YCONSERVATIVE
3.00	.80	1.00	1.00
4.00	.76	.96	.96
5.00	.73	.87	.86
6.00	.71	.75	.74
7.00	.68	.62	.61
8.00	.66	.52	.52
9.00	.64	.43	.43
10.00	.62	.35	.35
11.00	.60	.29	.28
12.00	.58	.24	.23
RANGE	ATM TRANS	XOPTIMISTIC	YOPTIMISTIC
3.00	.80	1.00	1.00
4.00	.76	1.00	.99
5.00	.73	.97	.96
6.00	.71	.92	.91
7.00	.68	.83	.82
8.00	.66	.73	.72
9.00	.64	.63	.63
10.00	.62	.55	.55
11.00	.60	.49	.48
12.00	.58	.42	.40



DETECTION PERFORMANCE TARGET DELTA TEMPERATURE IS 1.00 DEGREES C

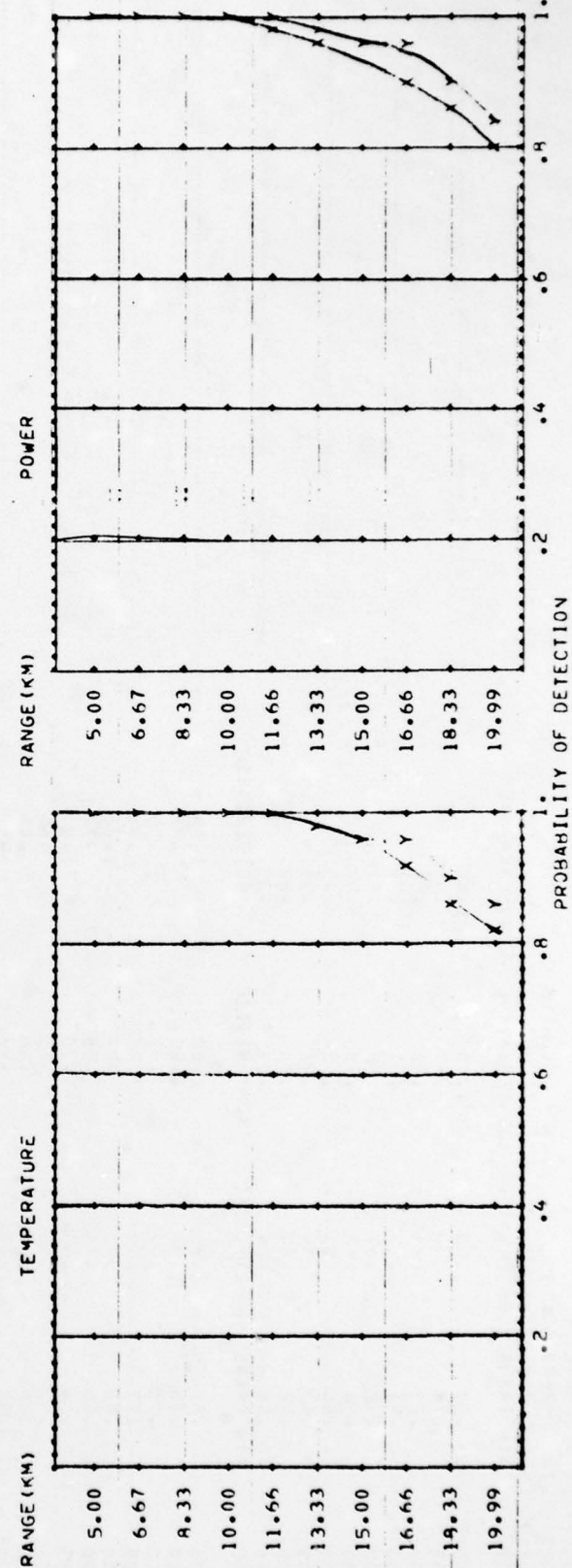
FOV PERFORMANCE TEMPERATURE DEPENDENT PERFORMANCE

RANGE	ATM TRANS	X DET T	Y DET T	45 DET T
5.00	.73	1.00	1.00	1.00
6.67	.59	1.00	1.00	1.00
8.33	.65	1.00	1.00	1.00
10.00	.62	1.00	1.00	1.00
11.66	.59	.99	1.00	.99
13.33	.56	.97	.98	.97
15.00	.53	.96	.97	.96
16.66	.51	.92	.95	.91
18.33	.49	.86	.91	.85
19.99	.47	.81	.86	.81

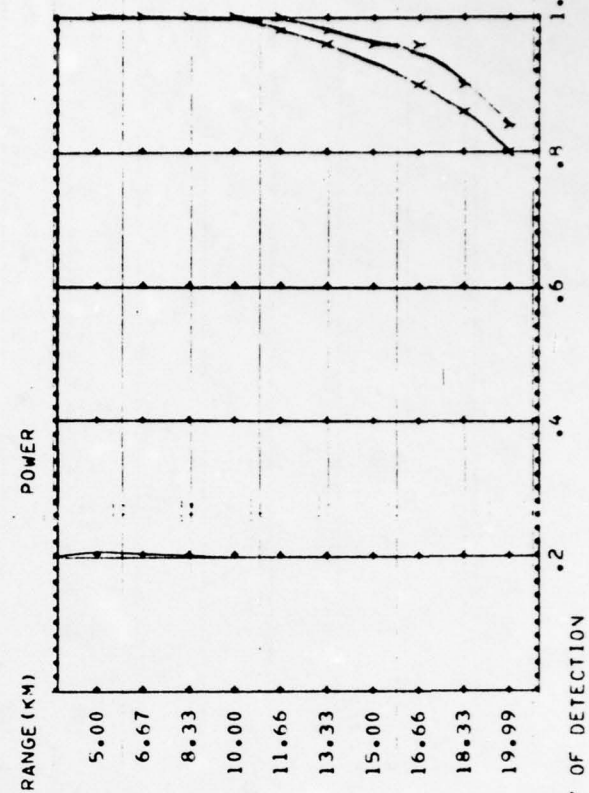
POWER DEPENDENT PERFORMANCE

RANGE	ATM TRANS	X DET P	Y DET P	45 DET P
5.00	.73	1.00	1.00	1.00
6.67	.69	1.00	1.00	1.00
8.33	.65	1.00	1.00	1.00
10.00	.62	1.00	1.00	1.00
11.66	.59	.99	1.00	.99
13.33	.56	.97	.98	.97
15.00	.53	.95	.96	.95
16.66	.51	.91	.95	.90
18.33	.49	.85	.90	.84
19.99	.47	.80	.85	.80

TEMPERATURE



POWER



DETECTION BASED ON MDT TARGET DELTA TEMPERATURE IS			1.00	DEGREES C
RANGE	ATM TRANS	S/N	PROBABILITY	
5.00	.733	824.658	1.000	1.000
6.57	.691	558.471	1.000	1.000
8.33	.653	406.622	1.000	1.000
10.00	.619	308.940	1.000	1.000
11.66	.588	242.374	1.000	1.000
13.33	.561	194.878	1.000	1.000
15.00	.535	159.601	1.000	1.000
16.66	.511	132.687	1.000	1.000
18.33	.489	111.594	1.000	1.000
19.99	.468	94.777	1.000	1.000

DETECTION PERFORMANCE
TARGET DELTA TEMPERATURE IS 1.00 DEGREES C

WFOV PERFORMANCE

TEMPERATURE DEPENDENT PERFORMANCE

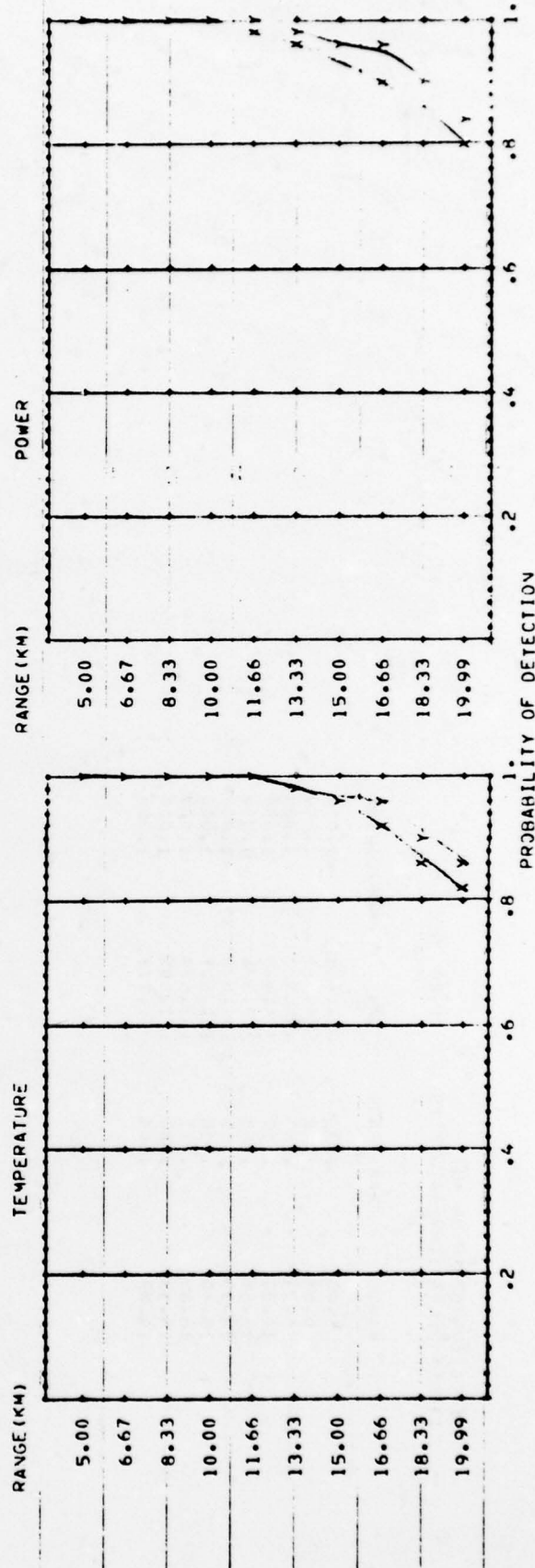
RANGE ATM TRANS X DET I Y DET I 45 DET I

5.00	.73	1.00	1.00	1.00
6.67	.69	1.00	1.00	1.00
8.33	.65	1.00	1.00	1.00
10.00	.62	1.00	1.00	1.00
11.66	.59	.99	1.00	.99
13.33	.56	.97	.98	.97
15.00	.53	.96	.97	.96
16.66	.51	.92	.95	.91
18.33	.49	.86	.91	.85
19.99	.47	.81	.86	.81

POWER DEPENDENT PERFORMANCE

RANGE ATM TRANS X DET P Y DET P 45 DET P

5.00	.73	1.00	1.00	1.00
6.67	.69	1.00	1.00	1.00
8.33	.65	1.00	1.00	1.00
10.00	.62	1.00	1.00	1.00
11.66	.59	.99	1.00	.99
13.33	.56	.97	.98	.97
15.00	.53	.95	.96	.95
16.66	.51	.91	.95	.90
18.33	.49	.85	.90	.84
19.99	.47	.80	.85	.80



APPENDIX B

CREATING THE SEARCH FIELD

OVERVIEW

The requirement was to produce a large number of search fields, each containing a target. The most feasible way to meet this requirement was by embedding the target in the scene. Computer embedding was preferred to photographic processing because of its high degree of control and flexibility.

Embedding is simply the substitution of selected picture elements in the scene by values corresponding to the target. This is a trivial computation for any computer. However, one must take care that the embedded target matches the overall scene with respect to size, intensity range, blur, orientation, and shadow position. Further, it is important to remove any artifacts associated with the embedding process such as light or dark edges surrounding the target.

We developed two sets of procedures for embedding targets. The first was described in the second interim report for the current contract.²

It was an interactive system using a plasma display of the targets and scenes and a large computer memory. However, the method was unsatisfactory, and a second--and much simpler--method was designed and implemented.

The embedding process starts with target and scene transparencies. These transparencies are analyzed by a scanning microdensitometer and put on magnetic tape.

CREATING THE TARGET FILE

The target transparency (25 x 50 mm) was scanned with a 100 mm analyzing aperture. The data on magnetic tape consist of 250 records with 500 bytes per record. Each byte contains the density (0.2 to 2.5) at each picture element.

The problem in digitizing the target is to separate the target from its background and outline it accurately. We ended up with a very simple approach. For each target a positive print was made. Then we cut out the target and its shadow with a scissors and pasted it on a white background. This was photographed and a transparency obtained. The transparency contained the target image on a clear background.

This target file was then collapsed into a 20 x 40 array by averaging over 12 x 12 pixel blocks. The 20 x 40 array was processed somewhat further. The target surround was made uniformly zero (i.e., completely transparent) and the gradient at the target border was sharpened because otherwise there would be occasional artificially light cells at the target border. We created a 78 target file in this way.

THE SCENE FILE

The scene file was constructed more directly. Each of the 24 scene transparencies (40 x 40 mm) was scanned with a 50 μ analyzing aperture. Thus, each scene was represented on magnetic tape as a file of 800 records, with 800 bytes per record.

EMBEDDING

The embedding process consists of selecting a scene, selecting a target, defining the target size and left-right orientation, placing it in a specified location in the scene. The output goes on magnetic tape and is then written out on film with a 50 μ writing aperture, resulting in a 40 x 40 mm transparency. The objective here is to choose the target size and location properly so that the final transparency is realistic in appearance.

APPENDIX C

SUBJECT INSTRUCTIONS

IN

NVL EXPERIMENT

AUGUST 1977

INSTRUCTIONS

We are trying to determine how long it takes people to find vehicles and other objects in different scenes.

To help us we carry out experiments like this one. We will show you a number of pictures and measure how long it takes you to find the targets.

Here are some of the scenes we are using.

EXAMPLES--PROJECTED ON SCREEN

Here are some of the targets.

EXAMPLES--8" x 10" HARD COPY

Here are some of the targets in scenes.

EXAMPLES--PROJECTED ON SCREEN

These have been made using a computer, and some may not be too realistic looking.

In the experiment the targets will always be in the regions outlined in red. You can see that all parts of an outlined region are about the same.

EXAMPLES--PROJECTED ON SCREEN

Sometimes the targets will be very easy to find and sometimes they will be very hard. Sometimes you won't be able to find the target at all, and at other times you will find it, but you won't be sure about what it is. And once in a while there may not be any targets at all.

Now we'll look at a tray full of slides. Try to find the target in 10 seconds. The targets may be very hard to find, especially at first. But that's the way the world is--a lot of the time it's hard to find what you are looking for. With some practice you will learn more about what the targets look like, and it will be a little easier.

After the slide has been on for 10 seconds, we'll stop and I will try to point out the target to you. Sometimes I won't be able to find it either.

And then we'll go on to the next slide. For now, you don't have to write anything down. Any questions? Let's begin.

-60 PRACTICE TRIALS-

(Now the experimenter went over each of the scenes, located the target, and explained why it was identified as such).

-CHANGE SLIDE TRAY-

Now let's practice some more trials. But this time we'll use the score sheets.

Your task is to find the target as quickly as you can. When you see something you think may be a target you don't have to be positive about it. If you are 80 percent sure, that's OK.

For each trial write down three things:

1. Time. 1, 2, 3, ... 10, or ---

Look at the clock in the room as you find the target and write down the time.

2. Target. Tank, Trucks, Jeep, Tractor, Van, Gun, Car, or X.
X means you think it is a target, but you can't tell at all what it is. Abbreviations are all right.
3. Location. Left (L), Middle (M), Right (R) in Region if it's in between, put down LM or MR.

Any questions? Let's start.

-10 PRACTICE TRIALS-

Let's look at these slides again.

-REPEAT TRIALS-

Any questions?

Remember, try to find the target as fast as you can. And you don't have to be positive when you find the target.

Try to spend as little time as possible figuring out what it is.

Now we will start the main session. Remember, you won't be able to find some of the targets. Just do the best you can. We'll stop for a moment after every 10 trials. If you have any questions at any time, please ask them.

-192 MAIN TRIALS-

Figure C-1 is an example of a typical test score sheet.

Tractor - ☒ A Jeep - ☒ J
 Truck - ☒ U Van - ☒ V
 Tank - ☒ T Car - ☒ C
 Gun - ☒ G

TL TM
LM M RM
BR BR

Name ALPH J NE S Date 9-12
 # of Targets Wing 33
 # of Loc Wing 35

	Time	Targ/Loc	Time	Targ/Loc	Time	Targ/Loc	Time	Targ/Loc
1	10	LM	61	7 LM	91	10 LM		
2	10	LM	62	10 LM	92	10 LM		
3	10	LM	63	10 LM	93	10 LM		
4	10	LM	64	3 LM	94	10 LM		
5	10	LM	65	10 LM	95	10 LM		
6	10	LM	66	10 LM	96	10 LM		
7	10	LM	67	10 LM				
8	10	LM	68	10 LM				
9	10	LM	69	10 LM				
10	10	LM	70	10 LM				
11	10	LM	71	10 LM				
12	10	LM	72	10 LM				
13	10	LM	73	10 LM				
14	10	LM	74	10 LM				
15	10	LM	75	10 LM				
16	10	LM	76	10 LM				
17	10	LM	77	10 LM				
18	10	LM	78	10 LM				
19	10	LM	79	10 LM				
20	10	LM	80	10 LM				
21	10	LM	81	10 LM				
22	10	LM	82	10 LM				
23	10	LM	83	10 LM				
24	10	LM	84	10 LM				
25	10	LM	85	10 LM				
26	10	LM	86	10 LM				
27	10	LM	87	10 LM				
28	10	LM	88	10 LM				
29	10	LM	89	10 LM				
30	10	LM	90	10 LM				

Figure C-1. Example of Typical Test Score Sheet